

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 04:04:07 ; Search time 7867.05 Seconds

10792.997 Million cell updates/sec

Title: US-09-893-033-6
 Perfect score: 1000

Sequence: 1 gccatagagccatcgccg.....cataagccaaattgtaatc 1959

Scoring table: IDENTITY_NUC

Searched: 3470272 seqs, 21671516995 residues

6940544

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: gēbā: *
2: gē hēg: *
3: gē jīn: *
4: gē ōm: *
5: gē ōv: *
6: gē pāt: *
7: gē pī: *
8: gē pī: *
9: gē pī: *
10: gē rō: *
11: gē sēs: *
12: gē sy: *
13: gē ūn: *
14: gē vī: *
15: em bā: *
16: em fun: *
17: em hūm: *
18: em jīn: *
19: em mū: *
20: em ōm: *
21: em ōr: *
22: em ōv: *
23: em pāt: *
24: em pī: *
25: em pī: *
26: em rō: *
27: em sēs: *
28: em ūn: *
29: em vī: *
30: em hēg hūm: *
31: em hēg jīn: *
32: em hēg ōther: *
33: em hēg mūs: *
34: em hēg pīn: *
35: em hēg rōd: *
36: em hēg mām: *
37: em hēg vīt: *
38: em sy: *
39: em hēg hūm: *
40: em hēg mūs: *
41: em hēg ōther: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result

No.	Score	Match Length	DB	ID	Description
1	1905.2	97.3	146174	1	D90910
2	924.6	47.2	1836	1	SPUNASK
3	878.4	44.8	300300	1	AP005373
4	849.2	43.3	342850	1	AP003597
5	683.4	34.9	1737	1	SPUNASL
6	672	34.0	297810	1	AP006579
7	647.4	33.0	349213	1	BX566693
8	638.4	32.6	2338	8	PSP224709
9	636	32.5	346683	1	BX572098
10	603.6	30.8	4099	8	EPECPTRA
11	603.6	30.8	191028	8	EPUS8804
12	592.8	30.3	149987	8	AB002583
13	592.6	30.3	121524	8	AF041468
14	584	29.8	164921	8	AF022186
15	566	28.9	300355	1	AE011162
16	562.6	28.7	349082	1	BX572091
17	47.6	21.3	300272	1	AE017213
18	416.6	21.3	1740	6	AR388434
19	415	21.2	10029	1	AE013966
20	415	21.2	201050	1	AJ414143
21	411.6	21.0	11759	1	AE011091
22	410.2	20.9	11318	1	AE013291
23	392	20.0	10028	1	AE012835
24	382.8	19.5	308050	1	SCO939124
25	382.6	19.5	349944	1	BX571871
26	382.6	19.5	349940	6	AX770909
27	381.6	19.5	44882	1	MLCB837
28	381.6	19.5	348950	1	MLEPPTN7
29	381	19.4	5806	1	AF117227
30	379.6	19.4	11367	1	AE004317
31	377.8	19.3	22348	1	AE008699
32	377.8	19.3	251050	1	AL627265
33	377.8	19.3	300169	1	AE016834
34	377	19.2	12167	1	AE005184
35	377	19.2	281530	1	AP002550
36	377	19.2	300409	1	AE016755
37	373.8	19.1	12065	1	AE004883
38	373.8	19.1	300933	1	AE004691
39	370.2	18.9	4655	1	AE015045
40	370.2	18.9	290029	1	AE016978
41	368.6	18.8	1734	6	EL1874
42	368.6	18.8	21757	1	AE000118
43	368.2	18.8	96109	6	AR048765
44	368.2	18.8	96109	6	AX067460
45	367.8	18.8	10029	1	AE006125
					D90910
					M75906
					SPUNASL
					AP005373
					AP003597
					SPUNASK
					AP006579
					BX566693
					PSP224709
					BX572098
					EPECPTRA
					EPUS8804
					AB002583
					AF041468
					AF022186
					AE011162
					BX572091
					AE017213
					AR388434
					AE013966
					AJ414143
					AE011091
					AE013291
					AE012835
					SCO939124
					BX571871
					AX770909
					MLCB837
					MLEPPTN7
					AF117227
					AE004317
					AE008699
					AL627265
					AE016834
					AE005184
					AP002550
					AE016755
					AE004883
					AE004691
					AE015045
					AE016978
					EL

ALIGNMENTS

RESULT 1					
LOCUS	D90910	146174 bp	DNA	linear	BCT 04-JUL-2001
DEFINITION	Synecocystis sp. PCC 6803 DNA, complete genome, section:12/27,				
ACCESSION	1430419-1576592.				
VERSION	D90910 AB001339 BA000022				
KEYWORDS	D90910.1 GI:1652956				
SOURCE	.				
ORGANISM	Synecocystis sp. PCC 6803				
REFERENCE	1. Bacteria; Cyanobacteria; Chroococcales; Synecocystis.				
AUTHORS	Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N.,				
TITLE	Sugita, M., and Tabata, S.				
	Sequence analysis of the genome of the unicellular cyanobacterium				

JOURNAL	Synechocystis sp strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome
MEDLINE	DNA Res. 2 (4), 153-166 (1995)
PUBMED	96127529
REFERENCE	8590279
AUTHORS	2 Kaneke,T., Sato,S., Kotani,H., Tanaka,A., Asamizu,E., Nakamura,Y., Miyajima,N., Hirasawa,M., Sugiyama,M., Sasamoto,S., Kimura,T., Hosouchi,T., Matsuno,A., Muraki,A., Nakazaki,N., Nanno,K., Okumura,S., Shimpou,S., Takeuchi,C., Wada,T., Matanabe,A., Yamada,M., Yasuda,M. and Tabata,S. Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions DNA Res. 3 (3), 109-136 (1996)
JOURNAL	97061201
MEDLINE	8905231
PUBMED	3 (bases 1 to 146174)
REFERENCE	Tabata,S.
AUTHORS	Direct Submission Submitted (28-JUN-1996) Satoshi Tabata, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; Yana 1532-3, Kisarazu, Chiba 292-0812, Japan E-mail: tabata@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/, Tel: 81-438-52-3933 (ex.2310), Fax: 81-438-52-3934 Potential protein coding regions were assigned on the basis of similarity search of the ORFs and Genemark analysis.
COMMENT	Location/Qualifiers
FEATURES	1..146174 /organism="Synechocystis sp. PCC 6803" /mol_type="genomic DNA" /strain="PCC6803" /db_xref="taxon:1148" /note="synonym: Synechocystis PCC6803"
SOURCE	1..486 /gene="apca" 1..486 /gene="apca" /note="ORF ID:s1r2067" /codon_start=1 /transl_table=11 /product="allophycocyanin a chain" /protein_id="BAI17874.1" /cd_xref="GI:1652957" /translation="MSIVTKSIVNADAEARKYISPEGLDIRIAFYTGGAARILAEITLTGSRETIYKQAGRLFGKRPDIVSPGNAVGEEMATCLRDMDYYLRVATYGVSGDVTPIEELIGLVGRMYRSLGTPLEAVAQSVEMKEVASGLMSSDDAAEASAYDFVIGXMS"
gene	582..1067 /gene="apcb" 582..1067 /gene="apcb" /note="ORF ID:s1r1986" /codon_start=1 /transl_table=11 /product="allophycocyanin b chain" /protein_id="BAI17875.1" /cd_xref="GI:1652958" /translation="MQDAITAVINSADVQGYLIDGAMDKLSYFASGELEVRASVYSANAATIIVEAKVKSLLYSDVTRPGANNWTTFRYAACIRDLDIYLRATYAMLGDGSLDERLYNGKETYNLSIGVPYISSVTOAIQAIKEVYASLVADAGKEMGVLDYICSEL"
CDS	1280..1483 /gene="apcc" 1280..1483 /gene="apcc" /note="ORF ID:s1r3183" /codon_start=1 /transl_table=11 /product="phycobilisome LC linker polypeptide" /protein_id="BAI17876.1" /cd_xref="GI:1652959" /translation="WMRFRTACVPSQGRIRIQRIORELONTTFYTKLVPYDNWMEQQRIIM"
gene	KMGKIVVEIATELAGPCGTNACIGA" complement(1578..2480) /gene="prna" complement(1578..2480) /gene="prna" /note="ORF ID:s111909" /codon_start=1 /transl_table=11 /product="ribosomal protein L11 methyltransferase; Prna" /protein_id="BAI17877.1" /cd_xref="GI:1652960" /translation="MAANSMWEIRILCHPLEETPAWRLEKFGCGLSTYEKAHSILLVGYLPQEKAREILDALALACEODDALFOYKPRFFAWOLIDEBSWSIKWHQPPVEDRPIIPAWIDPRENDRLILRDLPGAFTGTHTTOLCSLEBMRSEPKHCYTLADLGCGSLIGIGVAIVGAAYGVNDPLTVESARRHRLNOHPNLVINSESYELEOLAEPVVDGILCNIIAIEVVDLIPOPTPLVKRHGMALLSGINVEQSQAIALALENGMTVVATMKROQWCFCQARRREGD"
CDS	complement(2558..4222) /gene="sera" complement(2558..4222) /gene="sera" /note="ORF ID:s111908" /codon_start=1 /transl_table=11 /product="phosphoglycerate dehydrogenase" /protein_id="BAI17878.1" /cd_xref="GI:1652961" /translation="MLNALQLGSLZILLSPPAPALLIFRSFTMAKVLSIDPOVGIDLKOYAOVDVKTGLSEAITIDIVPEYAIMRSATKYTEKIIOKSQLXIIGRGVGVDNIDYPATRCGIOIYVNSPBGNTTAABEAHLAMMALAHRIPDANKSVKSKRRKOFICTEIVYKKTILGVGGIKTGSVAGYAKAMCKKLATYPTISOERADDIGTYVDILLFSHADFTLHIPTKPIETNALINAEILAMKFTARI INCSRGIIDEALVATALETDAI GGAALDVAFQBLGESRLREFSNILTPLHGASTEEAQVNVAVDAVBALDYVLLEGPASAVNIPELTDPVMEKRLPYLIKLAETIGTLVGOUGAGRIDRLTVLCLOGDLAEYTNQSP LVAVALKGLISCOLARERVNYAAIEAKERGI RVIETKDSVBDYSGSLHKCTGTWG EHSATGLINSGETIRITPDEFPINVPNNMMFLTRDMPEGITIKTGSLSGFNVMV ASMOYGKRIVRGDAIMALSDLDPLPDGLLSBITVNAIGIRAYTYKL"
gene	4483..5205 /gene="s1r1990" 4483..5205 /gene="s1r1990" /note="ORF ID:s1r1990 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAI17879.1" /cd_xref="GI:1652962" /translation="MIPLLTALAPGINFLRRLHRMSGLSINPATGQMOLAAGFLVSVPYFOAPIVROPFMISALTLPVNAIAVLMKSPROMMGDDLIFSMSGTAGAIFYWGFMPWPLMHLPIETALIGLPALMGKRGVCIHLYLISIMGTAATDLYFYLITGM DMRQVNVYVDEPFARFVFOALLAKITHPWCAANMILLALLILGVALOQPBPBGWA FEGAVLSTIVDALFWLGAMMA"
CDS	5227..6240 /gene="cyaA" 5227..6240 /gene="cyaA" /note="ORF ID:s1r1991" /codon_start=1 /transl_table=11 /product="adenylate cyclase" /protein_id="BAI17880.1" /cd_xref="GI:1652963" /translation="MDKPALPIYLVHPSYGVDTPEPLQGRSYTWTSQDNLDYIRON CTSRNHALIQATEEPTELLIDSRNGTFPNVARGVSVPILIODDKITFGKTBQFYS DONGTSPGLSPAPBWDTQTIVYHERRLITYLVADMENFTGMAQCYEBELISMILIGNWFQAGHILIREAGSWVDKXTIDAVNAIWPHGYNATPAEITIQILHAVRLOMWTKLNQ KYELPFPLRIGINTGYAMVNGTSGDHPTYTIGITVNAAPLESATQAHDLIAM SEKTSYVQDLPQWGAIVQCHTIELKGVTNITIIYGGFATIGTILDHTSHEDTTPAS EGP"
gene	6281..6745 /gene="s1r1992" 6281..6745 /gene="s1r1992"
CDS	/gene="s1r1992"

QY	426	GATTACTGAGAGGTGGGCGGTGCATGATTGTGACGATGCTTTCCAGGAAATTGACAT	485
Db	118808	GATTACTGGACAGGTGGGCGCGTGCATGATTGTGACGATGCTTTCCAGGAAATTGACAT	118867
QY	486	TTTTGGCAATCACTTACCGATTCGTTAAACACATCCATGTGGTACGTGTGCGCGGATAT	545
Db	118868	TTTTGGCAATCACTTACCGATGTTAAACACATCCATGTGTGGTACGTGTGCGCGGATAT	118927
QY	546	GCGTCGCATTGTTACTGAGGCTTTTCATCTTGCTAGACCCGATGTCGCGCGCGATTTT	605
Db	118928	GCGTCGCATTGTTACTGAGGCTTTTCATCTTGCTAGACCCGATGTCGCGCGCGATTTT	118987
QY	606	GATTCGATATTTCCAGAGATGTGGGCTTTGAAAGATGTGAGTACATTTCCCTCCAGCCCCG	665
Db	118988	GATTCGATATTTCCAGAGATGTGGGCTTTGAAAGATGTGAGTACATTTCCCTCCAGCCCCG	119047
QY	666	TGACGTTAAATCTACCGGCTTATCGCCCGACGGTTAAAGGTATCCCGCACAATTAATGC	725
Db	119048	TGACGTTAAATCTACCGGCTTATCGCCCGACGGTTAAAGGTATCCCGCACAATTAATGC	119107
QY	726	GCGATTGCAATTTGTTGAGCAGCGCCAGAAATCCCTGCTCTACGTAGGGGAGGGCGGAT	785
Db	119108	GCGATTGCAATTTGTTGAGCAGCGCCAGAAATCCCTGCTCTACGTAGGGGAGGGCGGAT	119167
QY	786	CGCGCGCAATGCGCCATGCGCCAGTGCAGAAATTTTGGGAAAGGTTCCAGTTGCCGGTAAC	845
Db	119168	CGCGCGCAATGCGCCATGCGCCAGTGCAGAAATTTTGGGAAAGGTTCCAGTTGCCGGTAAC	119227
QY	846	AACCACTCGATGAGAAATTGGGCTTTTGACGAAAAACCATCCCTTTGGTGAGTATGTT	905
Db	119228	AACCACTCGATGAGAAATTGGGCTTTTGACGAAAAACCATCCCTTTGGTGAGTATGTT	119287
QY	906	GGGTATGACATGCGCACCGCTATGCCAATTTTGGCGTACGGAATGTATTTGTGATGTC	965
Db	119288	GGGTATGACATGCGCACCGCTATGCCAATTTTGGCGTACGGAATGTATTTGTGATGTC	119347
QY	966	AGTGGGGGCGCGTTTTCGACGACCGGGTAACTGGCAATACCAATTTTGTAGCCGCGC	1025
Db	119348	AGTGGGGGCGCGTTTTCGACGACCGGGTAACTGGCAATACCAATTTTGTAGCCGCGC	119407
QY	1026	CAAAGTATTTACATTTGACATGATGACCCGCGGAGGTGGGAAAAACAGGGCTCCCGATGT	1085
Db	119408	CAAAGTATTTACATTTGACATGATGACCCGCGGAGGTGGGAAAAACAGGGCTCCCGATGT	119467
QY	1086	GCCCATTTGGGGGATGTACGCGCATGTTTGAACAAGTTTGGACGGGCGCGGAAATT	1145
Db	119468	GCCCATTTGGGGGATGTACGCGCATGTTTGAACAAGTTTGGACGGGCGCGGAAATT	119527
QY	1146	GGAATTACCCCAATCCCATGATACCAACCCAGGCAATGTTAATGCAATTGATCATTTGGCG	1205
Db	119528	GGAATTACCCCAATCCCATGATACCAACCCAGGCAATGTTAATGCAATTGATCATTTGGCG	119587
QY	1206	GACCGATTAACCCCTCCAGGTGCCCATGTAGAGGATCTATTTGCCCGCCAGAGGTAGT	1265
Db	119588	GACCGATTAACCCCTCCAGGTGCCCATGTAGAGGATCTATTTGCCCGCCAGAGGTAGT	119647
QY	1266	ACACGAAATTTGTGCGCAGGCGCCCGGATGCTTACTACACACCGATGTGGACAAACCA	1325
Db	119648	ACACGAAATTTGTGCGCAGGCGCCCGGATGCTTACTACACACCGATGTGGACAAACCA	119707
QY	1326	AATGTGGGGGGCCCAATTTTTTAACAATAGGGCCCCCGCGATGATTTCCAGTGTGGCTT	1385
Db	119708	AATGTGGGGGGCCCAATTTTTTAACAATAGGGCCCCCGCGATGATTTCCAGTGTGGCTT	119767
QY	1386	GGGTACGATGAGGCTTTTGTTTACCTGCGCGCATGAGGACCAAAATGGGAGTGGGGACG-	1444
Db	119768	GGGTACGATGAGGCTTTTGTTTACCTGCGCGCATGAGGACCAAAATGGGAGTGGGGACG	119827
QY	1445	AGGGGTCAATTTGATATAGTGGAGATGCCAGCTTCCAAATGAAATCTTCAGAAATCGGGAAC	1504
Db	119828	AGGGGTCAATTTGATATAGTGGAGATGCCAGCTTCCAAATGAAATCTTCAGAAATCGGGAAC	119887

QY 1505 CCTAGCCAGTACGATCCAGGTTAAACTATTCTCATTAACGGTTGGCAGGGAT 1564
 Db 119888 CCTAGCCAGTACGATCCAGGTTAAACTATTCTCATTAACGGTTGGCAGGGAT 119947
 QY 1565 GGTGCGTCACTGCGCAACAACTTTCTACGAAGAACGTTATTCGTTCTAAATGTC 1624
 Db 119948 GGTGCGTCACTGCGCAACAACTTTCTACGAAGAACGTTATTCGTTCTAAATGTC 120007
 QY 1625 GGGCATGCGACATTAATCTCTCTGGAAGCCATGAGTCAAGGGTATTAATGTCG 1684
 Db 120008 GGGCATGCGACATTAATCTCTCTGGAAGCCATGAGTCAAGGGTATTAATGTCG 120067
 QY 1685 CAAGCGGGAAGATTGTCGCCCGCGCATGCGGAAATGCTAGCCCAATGCTGTG 1744
 Db 120068 CAAGCGGGAAGATTGTCGCCCGCGCATGCGGAAATGCTAGCCCAATGCTGTG 120127
 QY 1745 GATGATGCTGTCGTCGCAAAAAGAGAACTGTTACCTTATGATGTCGCCCGCATGAG 1804
 Db 120128 GATGATGCTGTCGTCGCAAAAAGAGAACTGTTACCTTATGATGTCGCCCGCATGAG 120187
 QY 1805 TAATGCCCAATGCTAGGTTTACCGGAGTGCCTGTAACGACATGCTCCCGATG 1864
 Db 120188 TAATGCCCAATGCTAGGTTTACCGGAGTGCCTGTAACGACATGCTCCCGATG 120247
 QY 1865 GAGAGTCAACCTTCCCAACCCCAATTTTATCATACCCATGCTTCTGTTGTTGG 1924
 Db 120248 GAGAGTCAACCTTCCCAACCCCAATTTTATCATACCCATGCTTCTGTTGTTGG 120307
 QY 1925 AGCCAACTCTTACCCATTAAGCCAAATTTGAATTC 1959
 Db 120308 AGCCAACTCTTACCCATTAAGCCAAATTTGAATTC 120342
 RESULT 2
 SPUHASK
 DEFINITION SPUHASK 1836 bp DNA linear BCT 26-Apr-1993
 ACCESSION M75906
 VERSION M75906.1 GI:152904
 KEYWORDS acetylhydroxy acid synthetase.
 SOURCE Spirulina platensis
 ORGANISM Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.
 REFERENCE 1 (bases 1 to 1836)
 AUTHORS Milano, A., De Rossi, E., Zanaria, E., Barbierato, L., Ciferri, O. and Ricciardi, G.
 TITLE Molecular characterization of the genes encoding acetylhydroxy acid synthase in the cyanobacterium Spirulina platensis
 JOURNAL J. Gen. Microbiol. 138 (Pt 7), 1399-1408 (1992)
 MEDLINE 92381487
 PUBMED 1512571
 COMMENT Original source text: Spirulina platensis (strain C1) DNA.
 FEATURES
 source
 1..1836
 location/Qualifiers
 /organism="Spirulina platensis"
 /mol_type="genomic DNA"
 /strain="C1"
 /db_xref="taxon:1156"
 1..1836
 /gene="11vX"
 <1..>1836
 /gene="11vX"
 /EC_number="4.1.3.18"
 /codon_start=1
 /transl_table=1
 /product="acetylhydroxy acid synthase (AHAS)"
 /protein_id="AAA26594.1"
 /db_xref="GI:152905"
 /translation="VOLOTKIAARRATGAFRLIDSLKRGVQIIFGPGALIPIDE
 LYRAEADIOHILVRHOGASHADGVARATGTVGFCFGSGGADNLTGATAM
 DSIPIVITGQVAPALGTDAFODSDIFGLTPIVKSYVREBGMALVAEFHIA
 STGRPGVLIDVPDVGLEEDYIPVNGEVSILPGYPIYGVANROINOKILIEAE
 RPIWYVGGAISATNAHAIABLLFQIPVTTTLMGSGPDEKRPISVGLNMGHTAY

ORIGIN
 Query Match 47.2%; Score 924.6; DB 1; Length 1836;
 Best Local Similarity 69.7%; Pred. No. 1.9e-259;
 Matches 1265; Conservative 0; Mismatches 550; Indels 1; Gaps 1;
 ANFAVECDFLIAGAPFDDRVTKLDEFGSRKAVIHIDIDPAEVRKNTREVPVIGD
 VAOVILHLRRCREIGVNDQOTQSMERINRNPEDYPLVPSYDLSLQEVIAEL
 GMAPDGYTTTDPVGOHMAAOFLKNGROWISAGLGTMGVGIISAAGAKYALERSQ
 VCIAGDASVQMNIOELGTHQYGINVTVTLIINNQMOMVQMOQAFYGERYSASME
 ICMDFPEMLARSTYGVKMYVKSREDEHIAEMKLAIDYDPVIMDQVHTDENCIPNVAP
 GRSNAOMGLPERROLEKAVELIYSCNGKANNVSNPFCRGTKL"
 119 TTCTTAGCCACGCGCAACCGGGGCTTATCTCTGATGATAGCTTGAAGCCGATGGG 178
 Db 20 TTGCTGTCGCAAGGTGCAACAGGTGGCTTTCGCTCATGTATGCTTAAACGTCAATGGCG 79
 QY 179 TCAACACATTTTGGGCTATCCCGCGGGGCAATTTGGCCATCTATGATGAACGTAC 238
 Db 80 TACAGCACATTTTGGCTATCTGCTGATGAGCCATTCCTCCCATTTATATGATGAACGTATC 139
 QY 239 GCTTTAGACGGCGGGGAAATTTGAGCATTTTGTGCGCCGATGAACAGAGGCTTCC 298
 Db 140 GAGCGGAAGCCGAGGGCGATATACAGCATATCTGATTTGCGCATGAACAGGGGCGATCCC 199
 QY 299 ATGCGCGGATGGGTATGCGACAGCGACAGTAAAGTGGAGTTGTTGCGTATCATG 358
 Db 200 AGCGCGGATGGGTATGCGACAGCGACAGTAAAGTGGAGTTGTTGCGTATCATG 259
 QY 359 GACCAAGGGCGCATTAATCTGTGACCGCGCATTCGCCAATGCCATTTGAGCTCGTGC 418
 Db 260 GACCAAGGACCATTAATCTGTGACCGCGCATTCGCCAATGCCATTTGAGCTCGTGC 319
 QY 419 TGTGTGATTTACTGAGAGGTGGGCGCTGCTCATATTTGATAGCATGCTTTCCAGGAA 478
 Db 320 TGTGTATCATTAACCGACAGGTAGCGCGCGCCGATTTGATAGCATGCTTTCCAGGAT 379
 QY 479 TTGACATTTTGGCATACACTTACCGATGTTTAAACACTCTATGTGATGATGAGG 538
 Db 380 CAGATATTTTGGCATACACTTACCGATGTTTAAACACTCTATGTGATGATGAGG 439
 QY 539 CGGATATGCTGCGATTTGATGAGGCTTTTCACTTTGCTAGACACCGGTGTCGGGCG 598
 Db 440 GTGACATGCTGCGATTTGATGAGGCTTTTCACTTTGCTAGACACCGGTGTCGGGCG 499
 QY 599 CGGTTTGAATGATATTCGACAGATGAGGTGGGCTTGAAGAAATGATGATATTCCTG 658
 Db 500 CGGTTTGAATGATATTCGACAGATGAGGTGGGCTTGAAGAAATGATGATATTCCTG 559
 QY 659 ACCCGGTGACGTTATCTACCGGTTATCGCCCAAGGTTAAAGTATCCCGACAA 718
 Db 560 ATCAAGAGAGGTTCTTACCTGCTATCGCCCAAGGTTAAAGTATTCGCGCAA 619
 QY 719 TTAATGCGCATTTGATTTGAGAGCAGCGCAAAATCCCTGCTACAGAGGGGAG 778
 Db 620 TCAACCAACCATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
 QY 779 GGGCGATCGCGCCCAATGCCATGCCAGGTGACAGAAATTTGGGAAAGTTCCAGTTGC 838
 Db 680 GGGCGATTTCCGCGCAAGCCCATGCAAAATTTGGCAAGCTTTTTCAGATTTC 739
 QY 839 CGGTAAACAACCACTGATGGGAATTTGGGCTTTTGAAGAAACCATTCCTTTCGTTGG 898
 Db 740 CGGTAACTCAACATTAAGGGGAAAGATCTTTTGAAGAAAGATCCCTATCTGTTGG 799
 QY 899 GATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
 Db 800 GATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 859
 QY 959 TGAATGACATGAGGGGCGGTTTGAAGAGACCGGGTAACTGCGCAATTAAGCATTTGCTA 1018
 Db 860 TGAATGCTGATGAGGGGCTAGGTTTGAAGATCGGTTAAACCGGGTAACTTATGATTTGAT 919
 QY 1019 GCGCGGCGCAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1078

Db 920 CTAGGGCGAAGGTGATTCATATTTGACATTTGACCCCGGACGTAATGTTAAACCCGAGCC 979

Qy 1079 CCGATGTGCGCATTTGTTGGGGAGATGATACCGCATTTTGTGAACAGCTTTTGACAGGGGGCC 1138

Db 980 CCGAAGTGCATTTGTCGAGATGTTCCGACAGTGTGATACACCTGTTGGCCGCTGCC 1039

Qy 1139 GGGAAATGGATTACCCCAACCCATCCCATACCAACCGAGCATGTTAAATGTCATTGATC 1198

Db 1040 GAGAAATTTGAGATGTCGGGAGACGATTAATCAAACTCAATCGGTGTTAGCGGATTAAC 1099

Qy 1199 ATTGGCGACCGATTACCCCTCCAGTGGCCCATATGAGAGATCTATGCCCCCGAG 1258

Db 1100 GTTGGCCGGAAGATTATCTCTGGTGGTTCCTCTTATTTGATAGCCCTCTCTCCCAAG 1159

Qy 1259 AGGTAGTACACGAAATTTGTCGCCAGGCCCGCATGCTACTACACCAACCGATGGAGAC 1318

Db 1160 AGGTGATGCGAAGATTAGGAAAAATGGCACCGGATGGTACTATACACGAGATGGGTC 1219

Qy 1319 AACACCAATTTGGGGCGCCCGCTTTTGAACAATGGCCCCCGCGATGATTTCCAGTG 1378

Db 1220 AACACCAATTTGGGGCGCTCAATTTTGAATAATGGCCCCCGCTCAGTGAATCTACCG 1279

Qy 1379 CTGGCTTGGGTACGATTTGGGCTTTGTTACTGCGCCCATGGAGCCAAAGTGGAGTG 1438

Db 1280 CTGGGTTAGGAATCTATGGGCTATGGTATCCGCTCGGCTATGGAGCTAAAGTGGCTTGG 1339

Qy 1439 -GGACGAGCGGTCAATTGTCATGAGTGAATGCCAGCTTCCAAATGAATCTTGAGAAC 1497

Db 1340 ACGCATGCGAGGTGATTTGTATGCTGGGGATGCCAGTGTGAATGAACATCCAGAAC 1399

Qy 1498 TGGGAACCCCTAGCCCAAGTACGACATCCAGTTAAACTATTTCTCAATTAACGTTGGC 1557

Db 1400 TGGAAACATACGACATATGGCATTAATGTCAAGACGTAATTAATTAACGTTGGC 1459

Qy 1558 AGGGATGTTGCGTCAAGTGGCAACAACTTTCTACGAAAGCTATTTCTGCTTCTAACA 1617

Db 1460 AAGGATGTTGCGCGCAGTGGCAACAAAGCTTTTATGAGAGCTTATTCGGTTCTTACA 1519

Qy 1618 TGTCCAGGGCATTGCACGATTAATCTCTCTGTGAAAGCTATGGCATCAAGGTTATTA 1677

Db 1520 TGGAAATCGGAATGCGCAATTTTGAATGTTAGCACTTCCATGCGCTTAAAGGCATGG 1579

Qy 1678 CTGTGCGCAAGCGGAAATTTGGCCCGCGCATGCGCAAAATGCTAGCCCAATGGTC 1737

Db 1580 TCGTTAAGATGCTGATGACGATTAACACAGCATGACGAAATGTTGCTTATGATGCTC 1639

Qy 1738 CTGTGCTGATGATGTTGTTGTCATAAAAAGATGAAAACTGTTACCTATGATTTGCCCG 1797

Db 1640 CTGTGCTGATGATGTTGTTGTTACCAAGATGAAAAATTTGTTATCTATGTTAGACAG 1699

Qy 1798 GCATGATATGCCCAATGCTAGGTTTACCGGAATGCGGTACGAGCAATGTTCCCC 1857

Db 1700 GAGAGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1759

Qy 1858 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1917

Db 1760 AGTTAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1819

Qy 1918 GTTGTGAGACCAAACT 1933

Db 1820 AGTGTGACATAAAT 1835

RESULT 3
AP005373 300300 bp DNA linear BCT 17-AUG-2002
LOCUS Thermosynechococcus elongatus BP-1 DNA, complete genome, section
DEFINITION 5/9
ACCESSION AP005373 BA000039
VERSION AP005373.1 GI:22294885
KEYWORDS Thermosynechococcus elongatus BP-1
SOURCE

ORGANISM

Thermosynechococcus elongatus BP-1
Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.

REFERENCE

1 Nakamura, Y., Kaneko, T., Sato, S., Ikeuchi, M., Katoh, H., Sasaoka, S.,

Watanabe, A., Iriuguchi, M., Kawashima, K., Kimura, T., Kishida, Y.,

Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N.,

Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S.

Complete genome structure of the thermophilic cyanobacterium

Thermosynechococcus elongatus BP-1

DNA Res. (2002) In press

2 (bases 1 to 300300)

kaneko, T.

Direct Submission

Submitted (05-JUN-2002) Takakazu Kaneko, Kazusa DNA Research

Institute, The First Laboratory for Plant Gene Research, 2-6-7

Kazusa-Kametari, Kisarazu, Chiba 292-0812, Japan

(E-mail: kaneko@kazusa.or.jp)

URL: <http://www.kazusa.or.jp/cyano/Thermo/>

Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934)

Location/Qualifiers

1. 300300

/organism="Thermosynechococcus elongatus BP-1"

/mol_type="genomic DNA"

/strain="BP-1"

/db_xref="taxon:197221"

/note="BP-1"

/complement(293. 1519)

/gene="tll1162"

/complement(293. 1519)

/gene="tll1162"

/note="ORF ID: tll1162

putative transposase"

/codon_start=1

/transl_table=11

/protein_id="BAC08714.1"

/db_xref="GI:22294886"

/translation="MTAKSCVTHRIIAIMSGVEKAFSYRFPPTREOSILRRITGCV

RIYVNRALARTAEAWRKERLDYIOTLSLITQMKODILOFINEVSCVPLQALRHL

QSAFNSPFRGRATVPFKKRNQSGAEFTKSAFRMDKGYFLAKNCPINPMSRLP

DVPESTVITRLNPAQWYISLEFDPREITLQVPVSGLDVGISLITLSTGEKIA

NPGHFNRYKRLKQKRSLSRKQGSRNMDKALVAKIHQKISDKRKHDLQTLRL

IRENQIVIESLAVKIMVNRQIARISISAGWELVROLEYKAWGYRTLVKIDRFP

SSKRCGCGHIVEMLPISVEMDPCPKGAHHDINAAGNIIAVGHTVPCGAVPD

RHTSGGLRSTRSQK"

1689. 2072

/gene="tll1163"

/note="ORF ID: tll1163"

hypothetical protein"

/codon_start=1

/transl_table=11

/protein_id="BAC08715.1"

/db_xref="GI:22294887"

/translation="MANTPSRPEPLTLVBSPRSATAFVHLDSPVMTGLRRSLFTADGT

AVENQDLHTLALIELGYALASKRNOILVIOGHPGQAPLPFIEVSRORFVDTPT

LNANLATVORLLEPELFRGESLR"

complement(2082. 2774)

/gene="tll1164"

/complement(2082. 2774)

/gene="tll1164"

/note="ORF ID: tll1164"

unknown protein"

/codon_start=1

/transl_table=11

/protein_id="BAC08716.1"

/db_xref="GI:22294888"

/translation="MNTISPTRHYVRLLAEHPAIDSCIGCDRSRPHKINBEDLDRIYN

FOHLYPQNLQLEFLINLHQKIKANEPHQTLDQHTGRIFDLIGDLVSPNTTEL

LSFSEIANDNDKPAEPANOEATLTLPYEVLEFRITVEAVAYLQGMIVTNWMLTRP

RESWIAKIQVACFHNDPFRCEQLTTHAEATLQGMINHFIHRCQVLPILRLILEA

GIRSYLLYDLDVD"

3355. 3361

gene

gene
3861 .4454
/gene="clt1166"
3816 .4454
/gene="clt1166"
/note="ORF ID:clt1166"
/codon_start=1
/transl_table=11
/product="phosphopantetheine adenylyltransferase"
/translation="MLAYPGSFDDITTGHLDIRGARLISEYVAIAHPQKKALF
SVGGIKQVQAATSHLRKNRVDPDTFGLTVEARSGRAVLLLRGRVSDSEYELQMSH
TKNSLMPLEIETVFLTTSNEYSLTKEIARFGGNVRHLVPANVADLACPTQTR
IPSORPLE"
3816 .4454
/gene="clt1166"
3816 .4454
/gene="clt1166"
/note="ORF ID:clt1166"
/hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08719.1"
/db_xref="GI:22294890"
/translation="MLYSTBDISTTAGVNEALDPSAVILOLLOQLKEILLLEGT
KIPLGRKILDEQIQLQTLQIEEIPBSVEAAQILKROEIIKKQYRAQELIRRA
EORAAITELERIQOAELEAQIRIQOQVEEIMARCAIEMILRNTKEKLAHNR
OYTHCECOROEOEDAAVADRTLAEMERQLKEMIAVIQNGRQYLKQHPARQP"
complement(4548..5525)
/gene="clt1167"
complement(4548..5525)
/gene="clt1167"
/note="ORF ID:clt1167"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08719.1"
/db_xref="GI:22294891"
/translation="MNACQYMLSRVADAVMLNRYIERAENIARFVDVNLMLDLPT
SISLWMDPLVLTGDLSPFOEHTYGTAEENYIQLIPPTYNSLISGLRARNNAIS
IEVSSMDPEQVNAFYTMVBAAKAPGLLEIASLFEQKASHLFGVMSTSHNMS
AHNGQMGMLERADKSTRILDPVKYLLPSVEDGPTIDELGMALIKSAIAYEMVR
KRGSRITPAGVAFELIIDGERPRALRCLIOVEKSIYRITGTPLGSHQHOVEQOLR
LRSQLDYRITDELITISGMHREPLDIYIQOMNEVDYIOTPTFLPEVKR"
5565 .5897
/gene="ycf20"
5565 .5897
/gene="ycf20"
/note="ORF ID:clt1168"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC08720.1"
/db_xref="GI:22294892"
/translation="MORTRLDLRLCAGIQEOBKMPRRLATISIAFLFGVTLGIA
ISSAGQGLYDIIASSWVAIAEYISALPISDRKRLRQTLFGELNALKRGLLGF
IYAFLLGS"
6319 .7347
/gene="clt1169"
6319 .7347
/gene="clt1169"
/note="ORF ID:clt1169"
/codon_start=1
/transl_table=11
/product="pyruvate dehydrogenase E1 component, alpha
subunit"
/protein_id="BAC08721.1"
/db_xref="GI:22294893"
/translation="WVOESRLPSLSIPOTAIRBQGLMLYEDWMLGRFDEKCEWYI
BORMGFYLLNQGAVSGVTKAMRPDDYVCTSRDVAHLSAGIAPAREMALPDKR
ATGCGKRGSGSHLSAAXNPIFGSAPFAEAGIPATGGAPOTIKRIORMGAPKADQVY
ACFPEGGSNNQCFPECNLNALKPLILFVYENNNKMAIGAHBERASSETITTKYKAY
FQMWAEVDQMDVLAIVTVAEAIARARAGGPTILEALTYRFGHSLADEDELRSKE
EKEFMLEKDPIDIKQDAYLVEQELATGEDLRALIEQVQALVEDAVTFAEQSEPEEEL

gene	YVYFADP"	7406..7978	/gene="t1r1170"	
CDS		7406..7978	/gene="t1r1170"	
			/note="ORF ID:t1r1170	
			hypothetical protein"	
			/codon_start=1	
			/transl_table=11	
			/protein_id="BAC08722.1"	
			/db_xref="GI:22294894"	
			/translation="MLTLPHLFAMLDGLMDGIFSTOGIMILLMGYAGAMVFLSSA	
			PKRYITVMSEDLHARQFYEOELGLADIADVFHYVYNQQLGVASMDPLVYPALYS	
			SPRRPSPDGLVYOLKNTQTLAVISGASKGRNQRQRAVCPDRECLLELLRIQRQLR	
			YKMRSDRPINFLVKDYDGTVEMAEISARP"	
gene		7975..8901	/gene="t1r1171"	
CDS		7975..8901	/gene="t1r1171"	
			/note="ORF ID:t1r1171	
			hypothetical protein"	
			/codon_start=1	
			/transl_table=11	
			/protein_id="BAC08723.1"	
			/db_xref="GI:22294895"	
			/translation="MTNLQELPSRGHLITQGINPASCNLDQLTPELVLVFNQEDANT	
			LEALQAREALAIQTISATARLRQGRFLYGAGTSRLGLVDAECPPFCPLHIV	
			OGILAGDADLVRSSEGLDEYEDGAADVROTDADVIGISAGGTFFYVVGALDA	
			AGCVATLVEVACVPEQVORSADIDRLIVGPELILGSTRILKAGTAKVALNIISTG	
			VNVQLGKYGRNMDVAVNSRKLIDRARTILEDTGMDRAAALLERESGYKRALI	
			MMHTGLDAPAGALVDQNGQLHARRSHALDP"	
gene		complement (8898..9194)	/gene="t811172"	
CDS		complement (8898..9194)	/gene="t811172"	
			/note="ORF ID:t811172	
			unknown protein"	
			/codon_start=1	
			/transl_table=11	
			/protein_id="BAC08724.1"	
			/db_xref="GI:22294896"	
			/translation="WRLESTLTLAIFASTLTALPAFASDAIAGPIQVSPELRSALQAP	
			NIGASNMADQPVSMALLENRLLEISPVAREPREPVHQPEERPSGRFRKVDL"	
gene		9720..10169	/gene="t1r1173"	
CDS		9720..10169	/gene="t1r1173"	
			/gene="t1r1173"	
			/gene="t1r1173"	
Query Match 44.8%; Score 878.4; DB 1; Length 30030;				
Best Local Similarity 68.4%; Pred No.1.6e-245; Matches 1245; Conservative 0; Mismatches 572; Indels 4; Gaps 2;				
QY	135	AAACGGGAGCTTATATCCGATGATAGGCTGAAAGCCATGGGGTCAACATTTTGG	194	
Db	145700	AACTGGAGGACATCATTCGATGATACCTATGCGCGTACAGGTGTTAAGCATATTTTGG	145759	
QY	195	CTATCCCGCGGGGCAATTTTCCCATCTATGATGAACTGACCGCTTTAAGCGGGG	254	
Db	145760	ATATCTGTGGGGCCATTTTACCATTTATGATGAATTTTACCGTCTAGTCAAGGG	145819	
QY	255	GGAATTGAGCATATTTTGGTCCGCGATGAACAAGAGAGCTTCCCATCGCGGATGGGTA	314	
Db	145820	CATCATTCAGCATTTCTGTGTCCGCAAGAAAGGGGGCAGCCCATCTGTAGCGCTA	145879	145939
QY	315	TGCCAAGACCAAGGTAAAGTGGAGTTTGTTCGGTACATCTGGACCAAGGGCGCATTA	374	
Db	145880	CGCCCGTGCACAGGTCAAGTCAAGTGGGGGTGTGTTTTCGACTTCCGACCAAGGGGCAACAA	145939	145999
QY	375	CTTGTGACCGGATGGCAATGGCCATTTGGACTCGGTGCCATGNGNGATTAATCG	434	
Db	145940	CTTAGTTACAGGCATTCGCCACAGCCCAATGATTCATTCCTGCTGTGTGATCAGGG	145999	
QY	435	AGAGGTGGCCGTGCATGATTTGTAGGATGCTTTTCCAGGAATGACATTTTGGCAT	494	

Db 146000 TCAGGATACCTCCCTGCGCATTTGGACACCGATGCCCTTTCAAGAAACCGATATTTTGGGAT 146059
 Qy 495 CACCTTACCGATCGTTAAGACATCTCTATGTGGTACGTAGTGGGCGGATATGGCTCGAT 554
 Db 146060 TAGCGTCCCATTTGTTAAACACTCTTACGTGTGTGTGACGTAAAGGACATGGCACGGAT 146119
 Qy 555 TGTACTAGAGCTTTCCATCTTGTAGCAACGGTGTCCCGGCGGCTTTTGTATCATAT 614
 Db 146120 CATCTCTGAGCGCATCTACATTTGACACACATGTGTGACCCGGACCCGCTTGTGGATAT 146179
 Qy 615 TCCCAAGATGTGGGCTTAAAGAAATGTGATGACATTCCTCTCGACCCCGGTGACGTAA 674
 Db 146180 TCCCAAGATGTGGGCTTAAAGATGTGATGACATTCCTCTCGACCCCGGTGACCGGGGAGTGA 146239
 Qy 675 TCTACCGGCTTATCCGCCACCGTTAAGTATCCCGACAAATTAATGGCGATTTGCA 734
 Db 146240 GCTGCGGGCTATTCGCCCCACGATTTGCGGGAATGCCGTCAAAATTTGCCAAGGGTCCA 146299
 Qy 735 ATTGTGAGCAGCGCCAGAAATCCCTTGTCTTACGTAGGGGGAGGGCGCATGCCGCCA 794
 Db 146300 ATTGATTCGTAGGCGCCAGACGACGTTGTATGTGGGGGCGGCGCAATTAACGCGGG 146359
 Qy 795 TSCCATGCGCGAGTGTGACGAAATTTGCGGAAAGTTCCAGTTGCCGTAAACACACCT 854
 Db 146360 TGCTACAGAGAGGTCTCTCAAGTTAGCCGAGCTTTCCAGATTCAGTGAACACACACT 146419
 Qy 855 GATGGAATTTGGGCTTTTGAAGAAACCATCCCTTTGCGGGATATGGGTATGCA 914
 Db 146420 CATGGGGAAGGGGGCTTTTCTCGAATCCCATCCCTCTCTGTGGTATGTTGGGATGCA 146479
 Qy 915 TGGCCACCGCTATTCGCACTTTGCGGTACAGGAAATGTTGTTGATTTGAGTGGGGGCG 974
 Db 146480 TGGCACCGCTATTCGCACTTTGCGGTACGAAATGCAATCTCTGATTTGGGTGGGGCG 146539
 Qy 975 CGGTTTGAAGCAGCGGGTAACTGTGCAAACTAGACGAAATTTGTAGCCGCGCAAAATGAT 1034
 Db 146540 GCGCTTTGACATTCGCTACACGGTACAGTGTGACAGATTTGCTCCGCGCTAAGGTAT 146599
 Qy 1035 TACATTTGACATTCGACCGGGCGGAGGTGGGAAAAACAGGGCTCCCATTTGCCATGT 1094
 Db 146600 TCACATTTGACATTTGCTCTGAGGTGGGCAAAACCTGTCTCTGATTTGCCATGT 146659
 Qy 1095 GGGGATGTAGCCCATGTTTGAACAGCTTTTGAAGCGGGCGCGGCAATTTGATTAACC 1154
 Db 146660 GGGCAGCGTTAAACAGTGTCTGACAGCTCTCTAAGCAATTAAGACATGTGTCAAGC 146719
 Qy 1155 CACCATTCGCCATACCAACCCAGGCAATGTTAAATGCAATGATTCATTTGGCGACCGATTA 1214
 Db 146720 GGTTAACCTCAG--ACCCAAGCCTGTTAGAGCGGATTAATCTGAAAGCAGGATTA 146776
 Qy 1215 CCCCCCGAGGTGCCCACTATAGAGATATATTTCCCCCCAGAGGATAGAACAGAAAT 1274
 Db 146777 TCCCTCTATTTGTGCCCCCAACCAAGTGTGAATCTTACCCGAGAGGATGTTGCTTT 146836
 Qy 1275 TGGTGCAGAGGCCCGCGATGCTTACTACACACCGATTTGGGACAAACCAATGTGGGC 1334
 Db 146837 TGGCGGTACAGCCCGCGATGCTTATTAACCACTGATGTGGGTGACGACCAAGATGTGGC 146896
 Qy 1335 GGGCCAGTTTGAACATTTGGCCCCCGCGAGTGTTCAGTGTGTGCTTGGGTGATCAT 1394
 Db 146897 ACACACATTTCTGAAAAATGCGCTCGGGGTGTGATCTCCAGTGTGGGCTTTGGGACAT 146956
 Qy 1395 GGGCTTTGGTTACCTGCGCCATGAGCCAAAGTGGGAGTGGGGGA-CGAGCGCTCAT 1453
 Db 146957 GGGCTTTGGTTATGCGTGGCGCATGGGGGTCAAAAGTATGTTGCCGATCAACAGGTAT 147016
 Qy 1454 TTGCATCATGTAGATGCGAGCTTCCAAATGAATCTTACGAGAACTGGGAAACCTTAGCCCA 1513
 Db 147017 TTGTATATGTGGAGTGCAGTTTCCAGATGAATTTCCAAAGAGCTGGCTTACACTGGCTCA 147076
 Qy 1514 GTACACATCCAGTTAAACATTAATTTCTCATTAACGTTTGGCAGGGGATGTGTGCTCA 1573

Db 147077 GTACGATATTGCTGTCAAGACTGTCTATTATTAACAATTCTGCGACGGAATGTTGGCA 147136
 Qy 1574 GTGCAACAACAATTTTTCACAGAAAGATTAATTTGCTTCTTAACATGTCCAGGGCATGCC 1633
 Db 147137 GTGGCAGCAGGCTTTTATAGAGAGCGGTATTTCCATTCGAATATGGCCGAGGAATGCC 147196
 Qy 1634 AGACATTAATCTCTCTGTGAGACCTTATGSCATCAAGGATTAATCTGTGCGCAAGCGGA 1693
 Db 147197 CGATTTTGTAAATATGGCGAGGCTATGGGTTAAAGGGGCTACGGGTGAGCATGCGCA 147256
 Qy 1694 AGATTTGGGCCCGCGCATGCGCAATGTAGCCCAATGTGCTGTGTGATGTATGT 1753
 Db 147257 GGAATTTGATGATCCATTTGTCCAGAGGTTTGTGGCTTCAGATGGGCACTACTGCTGATGC 147316
 Qy 1754 GGTGTCMAAAAAGATGAAACTGTATACCTATGATTTGCCCGCGCATGATATGCCCA 1813
 Db 147317 CAAATGTACCCCGGATGAAATGTCTAACCAATGATGTGCCCCGCAAGCTTACTCTCA 147376
 Qy 1814 AATGCTAGTTTACCGGAATGCCGTACMGACATGTGTCCCGGATGTGTGAGTGC 1873
 Db 147377 AATGCTGAGTGTCCGGAACCAAGTGTCTTGAAGAGGCTGTGAACCTGCTTACTGTGC 147436
 Qy 1874 CCATGTGCAAAACCAAAATTTATCAACCATCGTTTCTGTCTGTGTGTGAGCAAACT 1933
 Db 147437 CAATTTGAGGCCAAGACGATTTCAACCATGCTTTTGTCCAGATGTGTGACGAAAGCT 147496
 Qy 1934 CTAAACCATTAAGCCAAATTTG 1954
 Db 147497 CTAGATTTCTTGACAAAGCTTG 147517

RESULT 4
 AP003597/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1. 342850
 /organism="Nostoc sp. PCC 7120"
 /mol_type="genomic DNA"
 /db_xref="taxon:103690"
 /note="synonym:Anabaena sp. PCC7120"
 complement(174..1478)
 /gene="all4593"
 complement(174..1478)
 /note="all4593"
 /gene="ORF ID:all4593
 hypothetical protein"

1
 Kaneko, T., Nakamura, Y., Wolk, C. P., Kuritz, T., Sasamoto, S.,
 Matanabe, A., Iriuchij, M., Ishikawa, A., Kawashima, K., Kimura, T.,
 Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A.,
 Nakazaki, N., Shimo, S., Sugimoto, M., Takazawa, M., Yamada, M.,
 Yasuda, M., and Tabata, S.
 Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120
 DNA Res. 8 (5), 205-213 (2001)
 URL: http://www.kazusa.or.jp/cyanobase/
 Tel: 81-438-52-3935 (ex. 2318), Fax: 81-438-52-3934)

```

/codon_start=1
/transl_table=1
/protein_id="BAB76292.1"
/db_xref="GI:17133730"
/translation="MDSKSTIQLVDELPSDNITVKNALDYVAPGQWQNLVGFNTR
AVGTETARVIOKTRDASTIYNDPOGQOSAVKLYOTIDKADTAMTALAKVCEK
IGFSLPSNITPKADYVOSTIDLVSKIAEIVPEKANGITPOPNQEFANSLANNYDA
SLIKFVALVCLDGLPLGPDPFSKIQGISAESNSITQNPVLAVNSLPSGSPIDK
LGFINGFNNAVQGMNNLVTKTGITPQSISSHGNFQIADNIDPVAALDDPTTFP
EHTGISOVARLLIKATLVKEIKOBTQPIQSSSAVSDTSQVAVSKIVMWE
BEDWQATIEKVDDOFFINYIGYSSHNWVGDDIRADYASDANGAVAVKATVC
MSDEETWTATIEKIQNGYFIHYGTVDSSYDEWVNSDDIS"
1680..1680
/gene="asr4594"
1680..1680
/gene="asr4594"
/nc="ORF_ID:asr4594
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76293.1"
/db_xref="GI:17133731"
/translation="MTALAPQEKSPROVLLYLGEDNYFVVEVPSLPQCISQKTRBEA
LTNIEVLQDKGELTPEYOFINI"
1877..2270
/gene="alr4595"
1877..2270
/gene="alr4595"
/nc="ORF_ID:alr4595
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76294.1"
/db_xref="GI:17133732"
/translation="MSKVDRLRLRFTFVYLGFTTHVQASAPFYEFINVDIA
VQPNGLVETQKNTFRDNLHLSRYTFLDNDRIEVSNGQLSTISTG
NHQLLSMEHQSPESHTEVLKRVIGLLVNNVAYWEALFSDRKAPLQAOVR
VEPEBELTNKIKIQSGVADIRVDTKTIEATKQPIEPGGLKQVYFDEAGKI
ETSGMSTQFFPERLVWILLGLVFGIPLLIIVSRNGTSSSYIGDSSIDYGGGG
DYGGDYGGDYGGGGGGGG"
3082..4473
/gene="alr4596"
3082..4473
/gene="alr4596"
/nc="ORF_ID:alr4596"
/codon_start=1
/transl_table=1
/product="malic enzyme"
/protein_id="BAB76295.1"
/db_xref="GI:17133733"
/translation="MADLTNPSFSVTLRQIPNRVGMASITQAIASSGNGIGQIDL
IEQROYSLRDLYDASTBEAETVOAVKAIPIKLVYDRITFNLHREGKLSISR
IPLKSVSLMAAYPGVAVCTAIADPSQVNTIKONTVAIVTDSAVLGNGNP
HALPVMGKMLPREFAGDAPPICATQDTEELIQAVKNIAVPFGVNLBIADPR
CFTEKRLRDLDPVPHDDQHTAIVTLAIPSLKLVOKSLADIRIVNGGAGV
AIALILKAKAEKIMWCDSKIIISTTTDLNEBQBPAYAKOCTLAGAVGAPVIGV
SAPGLTPBMYKSAKQIVPAMNP1PEIOPELISKQVAVIAGSDPQNNNVAA
FPVFGRLDQRAQITTTTMYLNASAIASLVNPSDLRHHIIPSVDERVTVAAVA
VQARBERGIARG"
5270..6853
/gene="alr4597"
5270..6853
/gene="alr4597"
/nc="ORF_ID:alr4597"
/codon_start=1
/transl_table=1
/product="type I restriction-modification system DNA
methylase"
/protein_id="BAB76296.1"
/db_xref="GI:17133734"
/translation="VTEKAKAKONNSSEFEOKMKADKLRKNTIDASYKHIVIGLIFL
KYISDAPEELHQKLAGEBDYAGANPEPDREYSANVFVPEARMSYLGQKQKQDPI
GKTYDLAMEAIEQENAKRLKGLIPKVVQCKQKLDQKSLGGLIDIGSITIGDARAQOD

```

```

VIGREYVFLQGFALAEGRKGGQFTYTPESIVKLLVENLBPYNGRFPDPCGGSGKQVQ
SEKPVKNQGRILDIDISIVGQESNETTYLCRMNLAIRGIDDSN1KNPESFLNDAHK
DLKADFLVIANPFPDSDWGBELRLNDGKMLDKDLYVPVPGNANFPMVSHFIVHLATGS
AGFVLSNGSLSSNTSGEDIDRKALVQKDLVDCIWLPTQLFNTGIPACIMFLSRYN
GNKNDRGVEVFLIDASELGVNVRSSAFTEASEISKIADTYHMKSGSGSYRDIKGF
CKSASIAIEKHNFVLTGRYVGI1PDEVEDGVSEFEKMSLTMLAGEMRGQGLDEE
IKKQLKXGFTVSDDELVP"
6858..7121
/gene="asr4598"
6858..7121
/gene="asr4598"
/nc="ORF_ID:asr4598
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76297.1"
/db_xref="GI:17133735"
/translation="MMMTQVILKLNPIVIEKLKHLAQSHQRTLEBEITSILEDVTE
TPITRSKRDSPGPFEOTCAGQCELLVRBPQEAQRBDPL"
7226..7804
/gene="alr4599"
7226..7804
/gene="alr4599"
/nc="ORF_ID:alr4599
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76298.1"
/db_xref="GI:17133736"
/translation="VIAYKGFPKLTPEEYIWEKQQLRREYINGEVYAMTGTQNH
GRIASNLI1FIVKGLRGGCQGVNSDCRIN1FEKDYVDPVSVTCBERDTA1QAO
YPCULIVEVLSPTASYDRGDKFRLVRRNPISQDYVLVDAEKIAD1LVRKNGRWELF
NYOLDN1KIOS1DIS1P1ESVYEDIVEE5V"
7855..8145
/gene="asr4600"
7855..8145
/gene="asr4600"
/nc="ORF_ID:asr4600
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76299.1"
/db_xref="GI:17133737"
/translation="MIFKEMMTQVILKQNLPIVLEKIKRIAQSHQRTLEBEITSILE
DVAKEKEEPKSEGEFMDMTLQFRERMOENITFDADPADIRDSVGRDVF"
8211..8573
/gene="alr4601"
8211..8573
/gene="alr4601"
/nc="ORF_ID:alr4601
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BAB76300.1"
/db_xref="GI:17133738"
/translation="MWKTLTQHOLETAGSVYMEELPFGCLRIYESQKRLLEVINQ
1PLKNTILNIDKLAQMRAPARLSKMGKTPAFIDQAIASINSLI1LVTVNVSDF
ESFNDLAVENFVNSGEG"
8577..9749
/gene="alr4602"
8577..9749
/gene="alr4602"
/nc="ORF_ID:alr4602"
/codon_start=1
/transl_table=1
/product="type I site-specific deoxyribonuclease chain S"
/protein_id="BAB76301.1"
/db_xref="GI:17133739"
/translation="MSEKKEKTLGEIADI1WVGSPTEGTCNNNGGULPLNGPTEFGD
RNPUPQOTIIPKXIAEGDLFCVRSSTTRMMADOKAYAGGIAISIRAKDILFQ
PYIRAIKEKSLAVATGTFPNIS1DHLNLI1VQPSKNIT1Y1SNLARIDEKI
YNSQNTLE1A1QTLFKHWFIDPEPNADGPKYKSSGAMVNSALGY1PEANSVGK
LGOYLN1KHGYAFRGEYITTEVTEKILLTPVNFK1GGGFNDSSKXYYSADYSNEVYL

```


LOCUS SPURASL 1737 bp DNA linear BCT 26-Apr-1993
 DEFINITION Spirulina platensis acetylhydroxy acid synthase (ilvY), 5' end.
 ACCESSION M75907
 VERSION M75907.1 GI:152906
 KEYWORDS acetylhydroxy acid synthetase.
 SOURCE Spirulina platensis
 ORGANISM Spirulina platensis
 Bacteria; Cyanobacteria; Oscillatoriales; Spirulina.
 1 (bases 1 to 1737)
 REFERENCES
 AUTHORS Milano, A., De Rossi, E., Zanaria, E., Barbierato, L., Ciferri, O. and Riccardi, G.
 TITLE Molecular characterization of the genes encoding acetylhydroxy acid synthase in the cyanobacterium Spirulina platensis
 JOURNAL J. Gen. Microbiol. 138 (Pt 7), 1399-1408 (1992)
 MEDLINE 92381487
 PUBMED 1512571
 COMMENT Original source text: Spirulina platensis (strain CI) DNA.
 FEATURES
 source
 Location/Qualifiers
 1..1737
 /organism="Spirulina platensis"
 /mol_type="genomic DNA"
 /strain="CI"
 /db_xref="taxon:1156"
 1..1737
 /gene="ilvY"
 1..>1737
 /gene="ilvY"
 /EC_number="4.1.3.18"
 /codon_start=1
 /product="acetylhydroxy acid synthase (AHAS)"
 /protein_id="AA26595.1"
 /db_xref="GI:152907"
 /transl_table=1
 /translation="MODQKQAVKRVTAAPALIDSLRRHGVQHIFGYPGSGNLPYDEI
 YRAQGEIKHYLVREHQAHAADGASTGKVGALVGLTGLATLAD
 SVPLATIGVPRALGTDAFOEIDIFGLPIVKSIVRSESLPIVAFHFLM
 SGPRGPLIDIPKDVNAOIDYIPEBGSRRVRYRPTBENRQI.NOLQISEATK
 PLIVGGGALMAHAIASELSERFQIPVSTLMGKRPHENPISLIGVLMGHGT
 AYANFAVMEIDFVAVGVRFDVAGIDQFAHSAKHIDIDPAEVGKRSDDVIV
 GDVAVQVIGMLQRTYHWEKLSRNKPNKGLDLQLSEPIELIVPHEBDGISPDGME
 LSHQCPDAFTTIDVGHQWMAQGFVQNGPRRMTSGGLGTMGLAAVGVKAHFHD
 TVTICSDGSFQMMOELGTIAQIGIKVYIINNGMLGVRQMHFMYNDRYEATNL
 EDGTEFARLADYGLAEMVVRKIKYORLPRALSHKGMILDVAVTDEDCYPMVA
 PGRHNSDMQLSS"

ORIGIN

Query Match 34.9%; Score 683.4; DB 1; Length 1737;
 Best Local Similarity 64.1%; Pred. No. 1.2e-188;
 Matches 1097; Conservative 0; Mismatches 601; Indels 13; Gaps 4;

QY 128 AAGCGCAACCGGGCTTATATCTGATGATAGCTGAACCGCAAGGGGTCAACACAC 187
 DB 26 AAGAGTAACAGAGACCTTGTGCTGATGATAGCTGCGCCGCAAGGTGACAGATA 85
 QY 188 TTTTGGCTATCCCGGGGCGGCAATTTTGCCTATGATGAAGTGTACCGCTTGAAG 247
 DB 86 TTTTGGCTATCCCGGGGCGGCAATTTTGCCTATGATGAAGTGTACCGCTTGAAG 145
 QY 248 CCGCGGGGGAATTTGAGCATTTTGGTGGCCATGAACAAGACCTCCCATGCCGCG 307
 DB 146 AAGCGGGGAATTAACATTATCTGTGCGGCAAGAACAGAGCGCCCAAGCTGCG 205
 QY 308 ATGGGATGCGAGACCAAGGTAAAGTGGAGTTTTCCTCGGATCTGACCAAGGG 367
 DB 206 ATGATATGCCCGAGAGTACCGGAAGGTGGGGATGTCTAGACACTTTCGACCGGAG 265
 QY 368 CGACTAATTTGATACCGGCAATTGCCAATGCCATTGGACTGGTGGCCATGGTGA 427
 DB 266 CCACCAATTTAGTACAGGTCTAGCCACTGCTTATCTCGATTCGTCGCCGTGTGA 325
 QY 428 TTACTGAGAGGTGGCGCGTGCATGATGGTGTAGCGATGCTTCCAGGAATTTGAATT 487
 DB 326 TTACCGACAGTCCCGCCGACGCGCTTGGGAACAGTGCCTTCAGGAATTTGATTTT 385

QY 488 TTGGCATCACCTTACCCGATCGTTAAGCACTCCCTATGCTGATAGTGGCGGATATAG 547
 DB 386 TTGGGATTAACCTTACCCGATCGTTAAGCACTCCCTATGCTGATAGTGGCGGATATAG 445
 QY 548 CTGCGATTGTTACTAGAGCTTTTCCATCTTGTAGACACCGTGTCCCGGCGGTTTGA 607
 DB 446 CGAGAAATTTGTGACAGGCAATTTCACTAGCCATGAGTGGACACCGGACCTGTGTTAA 505
 QY 608 TCGATATTTCCCAAGATGTGGCTTGAAGAAATGTAGTACATTTCCCTCGACCCCGGT 667
 DB 506 TTGACATTTCCCAAGATGTGGCTTGAAGAAATGTAGTACATTTCCCTCGACCCCGGT 565
 QY 668 AGTTAATCTACCGGGTATCGCCCAAGTAAAGTAAATCCCGCAATTAATATGCG 727
 DB 566 CAGTACAGCGGTGGGATATGTTCCCAAGACGGGGAATCTCGACAAATTAATATGCG 625
 QY 728 CATTGCAATTTGTGACAGAGCCAGAAATCCCTTGTCTTACGTAGGGGAGGCGATCG 787
 DB 626 CTTTACAGTTGATTTCCCAAGCTCTAAACCCCTCTATATGTTGGGGGAGGTCAATTA 685
 QY 788 CCGCAATGCGCATGCGCAGGTCAGAGAAATTTGGGAAAGGTTCCAGTTGCGGTTACAA 847
 DB 686 TGGCGGGGCTCATCTGAAATTCGCGAATTTGTGGAACGCTTCCAAATACCAATTACTA 745
 QY 848 CCACCTGATGGAATTTGGGCTTTTGAAGAAACATCCCTTTTCTCTCTCTCTCTCTCT 898
 DB 746 GCACTTTGATGGAAGAAAGTGGTGTGTAAGAAATCAACCCCTCTCTCTCTCTCTCTCT 805
 QY 898 GATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 958
 DB 806 GATGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
 QY 958 TGAATGAGTGGGCGCGCTTTGACAGACCGGCTTAACTGAGCAATTTGCTA 1018
 DB 866 TGATGCGGTGGGGTTCGCTTTGATGACCGGGTGGCGGATGATGATGATGATGATGATG 925
 QY 1018 GCGCGCCCAAGTAAATTTCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1078
 DB 926 ATAGGCGTGAAGTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 985
 QY 1078 CCGATGTCCTCATGTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1138
 DB 986 CTGATGTCCTCATGTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1044
 QY 1138 GGGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
 DB 1045 TATCATTTGGAACGGAATTTAAGCCGGAACCAACCCAGGATGCG--CACTGACCTTATC 1102
 QY 1198 ATTGCGGACCGATTAACCCCTCCAGGTGCGCCCACTATGAGATTAATTTCCGCCAGG 1258
 DB 1103 AATTGCGGAACCAATTAACCCCTCCAGGTGCGCCCACTATGAGATTAATTTCCGCCAG 1162
 QY 1258 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
 DB 1163 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1222
 QY 1318 AAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1378
 DB 1223 AGCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
 QY 1378 CTGCGTTGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1438
 DB 1283 GCGGTTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1342
 QY 1438 GGGAGCA--GCGGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1497
 DB 1343 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1402
 QY 1497 TGGGAACCTTACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557
 DB 1403 TGGGAACCTTACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1462

gene
CDS
DGLARTFAQGLEEEYAVAGIPIYNALITDEHSCQVAVDLITIRENYGKGLKAY
VGGDNVNVASHLITLCAKAYGSIIVATPEGFFPDPAVARSASEIAGRFGAVOLLRDPE
EAAERAHITLYDWTSMQGEAFETGHRQLPEQYINALLNCAAEAVIYHCLPAHNG
ZEITDEWVEGFRSRIMDEARSLHAQXAVIALAMGR
complement(945..2282)
/gene="g113102"
complement(945..2282)
/gene="g113102"
/note="ORF_ID:g113102"
/codon_start=1
/transl_table=1
/product="carboxyl-terminal protease"
/protein_id="BAC91043.1"
/db_xref="GI:35213673"
/translation="MKTPKRLPVGLACATVLSLSAFNMPSNFKDAPKEIIDEWV
QVNVREYDPTFNNNOMQETRRKLTADYSRDEAYKSTEDKGLGDPTFRMDPQ
YESMIDVAVNITKQAGTVALKRRRGKPLTDSVTRTELEKRVKSTLSEGGRRG
KAMDIDAVNITKQAGTVALKRRRGKPLTDSVTRTELEKRVKSTLSEGGRRG
YIRLSQFAVAAKMGNAVALKTRQADGFLIDRSPPGGLVASADIAILMTEGGIT

gene	AT1ADPQYBARVDVLGGQIGAMPVEQYVVAAGKFPFAERQPVAAENQTPVATP"
	complement (3391, .5712)
	/gene="g113103"
CDS	complement (3391, .5712)
	/gene="g113103"
	/note="ORF_ID:g113103"

```

/transl_table=11
/protein_id="BAC91044.1"

```

/translatom="MEADIPDNOCGDILAIYSRRIRIWEBSBDSRGKOTAFRNVLC
 HRPNTSLPRASLSEPGVAYPOSRGKGNHGRHRIYLSMSYRIGTGCTGTAFLVLA
 GQLGVGLASLAIIPRDSLAPSOQLAPLAPVGLAIVSEPTOPDADPAGAT
 IDLETVEGLSPRRSPVPTTREDIORKAANLAPALGAPVPEFANDARGDFHNG
 SYVRATLANOSMLINGRPIKGLNSHNGADYNSIPEVSGIRIELLPGAGTIVNS
 DYICGVATINITEEGRAVPTFSSASGEYGSIGYRYSYRGCTSRRLKINPGYSGATNN
 NPEVGSINNRCDDELADNADYSIYQSATITPDRNNLITDFLEPNBSKKEIJENB

FALDQWIDETRSFSPENLTKWQDZTRNLTSLNSAGNRWIPDIFPGLOQTDTATNGRI
 EADTQWIDETRSFSPENLTKWQDZTRNLTSLNSAGNRWIPDIFPGLOQTDTATNGRI
 EADTQWIDETRSFSPENLTKWQDZTRNLTSLNSAGNRWIPDIFPGLOQTDTATNGRI
 PDLEFLPDAAEFLFENRQDLEKPRGSGVYVNGAFDLEFADNRIGETVYGSLLQDINGNP
 LEADDTLQTRNINIGVNGTVSGAGIEMQDLEPMSSYLSLTKIKIKIEBGSLLGRLE
 PESVQWIDETRSFSPENLTKWQDZTRNLTSLNSAGNRWIPDIFPGLOQTDTATNGRI
 TVYVNLADVOYERNTVFSPGLTYRVQJQTSB"
 5711..6454
 gene="S1rj3.04"

```
/note="ORF ID:glr3104  
probable ABC transporter ATP-binding protein"  
/codon_start=1  
/transl_table=1  
/protein_id="BAC91045.1"  
/db_xref="GI:35213675"
```

gene
CDS
/cd_xref="01:35238675"
/translation="MLEAVDLTKNYGTAVPDLDGLNIRIAGEVFCLLGNGAKITTT
ISLGLFAPAGTAFTAKSDVQCHPLPETERFAVYIDQVMLYNIISGLNLEFSA
AGQGRYTEPELLAFPEQVGLPAESAGRGVYSGMKQKIAIILKKAEALLDDP
TSGDLPKSNFSFSLGRISROGVAVIMATHDLPRAKEAGTRVGINAGRLVTTARE
DLSYDLQERLTLESWGASAGGVREVR"
651..786
/gene="512105"
6451..7866

```

note="ORE_ID:GI173105
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC091046.1"
/db_xref="GI:35213676"
/cross_references="MIM:RAREY.EMWDRFRACATYVLGLVLSLMGMHYSAD
CA08AALCATTT001ADPPKPHSAHGVYAFKRLSPFDQGVDDYVGAAMLE

```

AAKNDPFRYPADATVAVORFGEITATVOLLPLVITLLGFGAPAGEEGTLEOL
 LSLGVPATLALGALGMVALGILLVPAITLLGVALALAGGGIAGLAPLALISA
 AYLGYLLFTVGYCLAVSAMAKSARIALVALGFWIYOGIALAPRYDLARGLYPUSA
 LTTETVATDMAGGIDGSHNSQDARPALEKRVIOQYASKEVDLPVNFICITAGESE
 YGYRVDRRFGELMRPEEONAVHOGFVAPILAVRSMSJSGDVFQHRFPATA
 EOYRGRIKTMENAMTENKDDGYRATADYRSVDFQYTPPLGWTAAQOSJLLGLG
 FWIYLAANAFAAARLRYTG"
 7866..9311
 /gene="g1r3106"
 /note="ORF ID:g1r3106
 unknown protein"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC91047.1"
 /db_xref="GI:35213677"
 /translation="MTASITKHEMRNLMARVIMVYGGFLPALLIVAGLNGTGMVOFO
 RQILAQAQSEITRLDPLQSIADIESGKRFQDAPADPROPTVGNKGRVATLT
 RPLAPLAVQSDLYPYFVSTLNKQTFIÖNNELNPTNLGRDPLAFVLYLLPL
 FILAGYNLSGERDEBETLAMLISOPVGLMVLVAGGIVIALAVLSIAGVL
 GGADPAEPGRRLRLGLMVLVILIGVMEFLAVAVNAGRTPTNAVLVGLMLLV
 IVPVVAVASTSPVSRVYLIOAVREATNANAGSRLIARYEDHDELPKDSRV
 DQGFALRSTAVRETVDKATAPLERYDROLAQQALVRYVYLSALIVAOEVNDLA
 GTGLGRHRHOVDRHRTWQGYFVPRIRQOTVTAELGRLPKRFREBELTGDVAG
 RVLPGVLAPLVLVGLTLPGLRRIPVA"
 complement (9377..10342)
 /gene="g113107"
 complement (9377..10342)
 /gene="g113107"
 /note="ORF ID:g113107
 hypothetical protein"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC91048.1"
 /db_xref="GI:35213678"
 /translation="MRIIVVALLMMGLATPAPAPLKIQYSAMPQMPMKVABAKGL
 PKKNGVQWVMEFVADSTITALVAGLNDLCOTLDTTAPAAAGAKLVAVLNTNSA
 GNOIILYKRIASIKDKKKKALGELCTDHYLLIGKEGLTQDLOPTPLTGA
 AARPAGRLDVGVPAPFTKALSTGKVLFSKKEPQGISDHLVLRVAKRGE
 IERIKVTWFTDLYIKNPESLQIMAKQGSTVAEKSDGSTRIFNLÖNLVTFAP
 ETQDILRIYOSKIIDFLVKSGLAKTKPDTAVLPEDKFRALKP"
 10476..10730
 /gene="g3r3108"
 10476..10730
 /gene="g3r3108"
 /note="ORF ID:g3r3108
 unknown protein"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC91049.1"
 /db_xref="GI:35213679"
 /translation="WPIHADLTRHFEETPLPSLPZHRAARTLHÖKREKLDALERS
 TGMFTAQETARACAKELVDATETREAKIVLKAQPO"
 complement (10782..11156)
 /gene="g113109"
 complement (10782..11156)
 /gene="g113109"
 /note="ORF ID:g113109
 hypothetical protein"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC91050.1"
 /db_xref="GI:35213680"
 /translation="MAKNKRSOTPROIPGLPVPGNRVRLDPOVLDAGRVAVI
 DCSASRRIMTLAGVAPDARNFEINNRKGVVVCVMWLSGRSLVAAYRLAKGCT
 AYDLKGLAAWKTRNGYPTIRAL"
 11432..11905
 /gene="g1r3110"
 11432..11905
 /gene="g1r3110"
 /note="ORF ID:g1r3110
 unknown protein"

Query Match 34.3%; Score 672; DB 1; Length 297800;
 Best Local Similarity 62.4%; Pred. No. 5,88-185;
 Matches 1127; Conservative 0; Mismatches 656; Indels 22; Gaps 4;
 /codon_start=1
 /transl_table=1
 /protein_id="BAC91051.1"
 /db_xref="GI:35213681"
 136 ACCGGGGCTTATATCCGATGATAGCTGATAAGCTGAGGAGGCAACATTTTGGC 195
 181531 ACCGAGGCTTCGCGCTGCTGATAGTGTGCTGCAAGGAGGCTCCGACATCTTCGT 181590
 196 TATCCCGAGGGGCAATTTTGCCCATATGATGAATGTAACGCTTTGAAGCGCGGG 255
 181591 TACCCGGTGGGCGCATTTTGCCCGGTCTACAGACAGATTATCAAGGCGGAGCGG 181650
 236 GAATTTGAGCATTTTGTGTGCTGCGCCATGAAACAAGACCTTCCATGGCGGAGAT 315
 181651 CGGTTGAGCATTTTGTGTGCTGCGCCATGAAACAAGACCTTCCATGGCGGAGAT 181710
 316 GCCAAGCCAGAGTAAAGTGGAGTTGTTTGGTATCATCTGACACCAAGGCGCATAC 375
 181711 GCCCGGCGAGCCGGCGGGGTGGGTGTGTGTGCACTCCGCTCCGGGGCGAGAT 181770
 376 TTGGTGAACCGCATTTGCCAATGCCCATTTGACCTCGGTCCTATGTTGATTTACGA 435
 181771 CTGGTGAACCGCATTTGCCAATGCCCATTTGACCTCGGTCCTATGTTGATTTACGA 181830
 436 GAGGTGGCGCGCATGATTTGTAGCATGCTTTCCAGAAATTTGACATTTTGGCATC 495
 181831 CAGGTGGCGCGCGCATGATTTGTAGCATGCTTTCCAGAAATTTGACATTTTGGCATC 181890
 496 ACCTTACCGATGTTAGCACTCTGTATGTGTAGTGTAGTGTGTGTGTGTGTGTGT 555
 181891 ACCGTGCGCATGTTAGCACTCTGTATGTGTAGTGTGTGTGTGTGTGTGTGTGTGT 181950
 556 GTTACTGAGGCTTCCATCTTGTGTAGCAACGGGTGTGTGTGTGTGTGTGTGTGTGT 615
 181951 TTACGCAAGCTTCCATATGCAAGCAACGGGTGTGTGTGTGTGTGTGTGTGTGT 182010
 616 CCCAAGATGTGGGCTTAAAGAAATGTAGTACTTCCCTGACCCCGGTGACGTTAAT 675
 182011 CCCAAGATGTGGGCTTAAAGAAATGTAGTACTTCCCTGACCCCGGTGACGTTAAT 182067
 676 CTACCGGCTTATCGCCCAACGCTTAAAGTAAATCCCGCAAAATTAATGGCGCATTCGA 735
 182068 ATTCCCGGCTTATCGCCCAACGCTTAAAGTAAATCCCGCAAAATTAATGGCGCATTCGA 182127
 736 TTGTTGAGCAGGCGCAAAATCCCTGCTACGTAGAGGAGGCGCATTCGCCCAAT 795
 182128 CTGATCCCTGAGAGGAGCGCGCTCTGTTATGTGGGTGGCGCGCATTCGCCCGGG 182187
 796 GCCATGCGCGAGTGCAGAAATTTGCGGAAAGTTCCAGTTGCCGTTAAACAACCTCTG 855
 182188 GCGCAGCGCGAGTGCAGAAATTTGCGGAAAGTTCCAGTTGCCGTTAAACAACCTCTG 182247
 856 ATGGGAATTTGGGCTTTTGAAGAAACCATCCCTTCCGAGGATGTGTGTGTGTGTGTGT 915
 182248 ATGGGCAAGGGGCTTCCCGAAGATCATCCGTGCGCGTGGGCAATGCTCGGATGAC 182307
 916 GCGCAGCGCTAGCCCACTTTCGCTGACGCAATGTATTTGTATGATGCAATGTGGGCGC 975
 182308 GCGCAGCGCTAGCCCACTTTCGCTGACGCAATGTATTTGTATGATGCAATGTGGGCGC 182367
 976 CGTTTCGAGCAGCGGGTAACTGGCAAACTGAGCAATTTGTAGCGCGCAAGTATTT 1035
 182368 CGTTTCGAGCAGCGGGTAACTGGCAAACTGAGCAATTTGTAGCGCGCAAGTATTT 182427
 1036 CACATTTGACATGACCCGCGGAGGTGGGAAAAAACAAGGCTCCCGATGTGCTCCATTTGTG 1095
 182428 CACATTTGACATGACCCGCGGAGGTGGGAAAAAACAAGGCTCCCGATGTGCTCCATTTGTG 182487

QY 1096 GGGAGTGTACGCCATGTTTGTAGACAGCTTTGACGCGGCGCCGGAAATGATTAATCCCC 1155
 Db 182488 GCGGCGGTCAAACTGGTGTGACGATCTGATTAATAAGCTCATGAAATTT----- 182537
 QY 1156 ACCCATCCCATACACACCGGATGTTAAATGCGATTTGATTCATTTGGCGGACCGGATTAC 1215
 Db 182558 -----TCAGCCGCGGACCCGGCGTGGCTGGAGCCCATGACGCTGGAAGGCGAGTAT 182592
 QY 1216 CCCCCTCCAGGTGCCCCACTATGAGATTAATTTGCCCCCGACGAGATGATACAGAAATT 1275
 Db 182593 CCGCTGAGAGTCCCGAGGCGATGACACCGGCTGTATCCACAGGAAGTATGTTGCCCTTT 182652
 QY 1276 GGTCCCGCAGGCCCCCGGATGCTTACTACACACCGATGTGGGACAAACCAATGTGGCG 1335
 Db 182653 CGGAGCTGGCTACCGAGCGCTACTACACACCGAGTGGCGACACCAAGATGTGGCG 182712
 QY 1336 GCCCAGTTTTTGAACATGCGCCCGCGGATGATTTTCAGTGTGCTTGGGTAACGATG 1395
 Db 182713 GGGCAGTTTCTCAAAAAGGCGCCCGCGCTGGATTTCTCAGCGGCGCTCGGACGATG 182772
 QY 1396 GCGTTTGGTTTACCTGCGGCCATGAGGACCAAAGTGGAGTGGGGACG-AGCGTCAATT 1454
 Db 182773 GCGTTCGCGCTTCCGCGCGCCATCGCGGTGACGAGTGGCACTGCTGACGAAGCGTCGTG 182832
 QY 1455 TGCATCAGTGGAGATGCCAGCTTCGAAATGATCTTCAGGAACCTGGGAACCTAGCCGAG 1514
 Db 182833 TGCATCTCCGCGGACGCGGCTTTGATGATTAACCTACAGGAATGGGACCGCGTGCAG 182892
 QY 1515 TACGACATCCAGTTTAACTATTAATTCATTAACGTTGGCAGGGGATGTTGTCAG 1574
 Db 182893 TACGACCTGCGCGTCAAGGCGCGCATCGTCAACAGCGCTGGCAGGGCATGTCGCGCAG 182952
 QY 1575 TGGCAACAACCTTTTACGAGGAAGGATTAATCTGCTTCTAATCATGTCCAGGGCATGCCA 1634
 Db 182953 TGGCAAGAATTTTCTACGACCGCGCTTCCCATCCACATGAAGCCCTCGATGCC 183012
 QY 1635 GACATTAATCTCTCTGTGAAGCTTATGACATCAAGGATTAATCTGTCCGACAGCGGAA 1694
 Db 183013 GACTTCGTCAAACTGCTGAAGCGTTGGGGGCGAGGGGACGTGTCATGATCAACCCGAC 183072
 QY 1695 GATTGGCGCCCGGCGATGCGGCAATGCTAAGCCCAAAATGCTCTGTGTGATGATGTG 1754
 Db 183073 CAGCTCAAGAAGGAGTGAAGCGCTGATGAGCGCGGACCGCGCTGATTTGCGACTTT 183132
 QY 1755 GTGTCAAAAAAATGATAAATGTTTACCTTATGATTGCCCCCGGATGATTAATGCCCAA 1814
 Db 183133 CATGTGCTGCCGAGCGCACTGCTACCGGATGTGCCCCCGGCGACAGCCACACAG 183192
 QY 1815 ATGTAGGTTTACCGGAAGTCCGGTACNGACAATGTTCCCGGATGTTGAGTGCAC 1874
 Db 183193 ATGTGCGGGCTGCC---CAGGCGCTTGCCTGATGCGAGCGCTGAACTCATCTACGCC 183249
 QY 1875 CATTCGCAAAACCCAAATTTTATCATCACCATGTTCTCTGTTGTGGAGCCAACTC 1934
 Db 183250 AACTCGGCAACCAAAATACACCGCAACGCGCTTGCCTCCGAGTGGGACCAAGCTT 183309
 QY 1935 TAACC 1939
 Db 183310 TAGCC 183314

RESULT 7
 BX569693/c BX569693 349213 bp DNA circular BCT 23-SEP-2003
 DEFINITION *Synechococcus* sp. WH8102 complete genome; segment 5/7.
 ACCESSION BX569693 BX548020
 VERSION BX569693.1 GI:33638930
 KEYWORDS complete genomes: WH 8102
 SOURCE *Synechococcus* sp. WH 8102
 ORGANISM *Synechococcus* sp. WH 8102
 Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.
 REFERENCE 1 (bases 1 to 349213)
 AUTHORS Patenik,B., Brahamsha,B., Larimer,F., Land,M., Hauser,L., Chain,P.,
 Larmerdin,J., Regala,W., Allen,E.A., McCarren,J., Paulsen,I.,
 Dufresne,A., Partensky,F., Webb,E. and Waterbury,J.
 The genome of a motile marine *Synechococcus*
 Nature 424 (6952), 1037-1042 (2003)
 JOURNAL MEDLINE 22825697
 PUBMED 12917641
 REFERENCE 2 (bases 1 to 349213)
 AUTHORS Larimer,F. and Patenik,B.
 CONSRM *Synechococcus* genome consortium
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-2003) Larimer, F., Submitted on behalf of the
Synechococcus genome consortium, the DOE Joint Genome Institute,
 Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA
 94598, USA, and the Genome Analysis Group, Oak Ridge National
 Laboratory, 1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
 larimerf@ornl.gov

FEATURES
 source location/Qualifiers
 1..349213
 /organism="Synechococcus sp. WH 8102"
 /mol_type="genomic DNA"
 /strain="WH8102"
 /db_xref="taxon:8458"
 /complement(52..1029)
 /locus_tag="SYNW1422"
 /complement(52..1029)
 /locus_tag="SYNW1422"
 /note="Similar to selenophosphate synthetase"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAE07937.1"
 /db_xref="GI:33638931"
 /translation="MSFTLMDGLRLHSLGKLAKRDIOPAGNFCQPPQLGAGMLG
 DDAALLPQNGQLLACGEOPLVEDEDFWAGSGVLYNLSDIAMGRPIALVNSV
 WSGADALPQLLEGMKFAFCEGFGVWQSGPYTLVSVLGVAGPILASAA
 APDESLMMVNNKGAFYRHPMDATHTPPALREHSLTLPALAEGLVRAKIDSM
 GGLTSTWMPAEACGLESLDIDAVDRPVGVEQMLSCFSPRGVILADPRTSTLV
 RMQGPDLICRIGRFSTGBCRVLIQISGASHCLMEVASGLTGRGCG"
 /complement(1026..1697)
 /locus_tag="SYNW1423"
 /complement(1026..1697)
 /locus_tag="SYNW1423"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAE07938.1"
 /db_xref="GI:33638932"
 /translation="WVFCIDPSRSIGRSISSAPSLFTPSVRAIGIDDDPRLSPTA
 SSRPSFHLIRPDSPLIAGWLSRSIFCEGHVPEQSDRLDRACIALIHHQD
 PBDKGPSSGVGVYRIYETRPRLMYGRGLGVHSDFRHNOIGKLIRKATITANGW
 GCDRFAMVQIQWRFRLHMASIEBELRIGRIHLMQADNYTAPSRERRPSCLL
 PSLAA"
 /complement(1700..2791)
 /locus_tag="SYNW1424"
 /complement(1700..2791)
 /locus_tag="SYNW1424"
 /note="Biotin synthase-like domain containing protein"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAE07939.1"
 /db_xref="GI:33638933"
 /translation="MSEIGRLITLLOVKVRAPELNGRGREGAGSPDRALNDGT
 TTVVPRYNVSSSAVSLAARDGLITRFPQDAPLVTTTDEPFLYGLSTREGIPYR
 STALHSRVLAITLLQCTIRDRSQSCQFALIOSLEDGATVIRKPEQVAAEA
 AVRLDGVKQLVMTGTPNSDRGRALMETAAVRRVLDLPQGGCEPDDIITWERM
 KAAGIDSLGMLHEVLEVRRIILPGKSLISLERYAAFDADVAVAFGGEVSTYLLAG
 LQDSREALDCCRLILEGVYFPVPIISGTPLEHHAPDTPAMVDVYKGVALLH
 OGLIREAMAGCACGCSALSTPEQAA"
 /complement(2784..3800)

gene
 CDS
 gene

```
gene
CDS
    /locus_tag="SYNM1425"
    /complement(12784..3800)
    /locus_tag="SYNM1425"
    /EC_number="3.5.5.7"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="probable nitrate reductase"
    /protein_id="CAE07940.1"
    /db_xref="GI:33638934"
    /translation="MTVTVAQAQRPVLFSLDGLQKYLDAEAAGVGLVLP
    ETRPYRPSFVPPVPMGRSHLALYDAVYVPGPTDAVAAARQGMQVILGNE
    RDGGLTNTQLTNSCGEIALKREKTPITHERMWQGGSGLSVSTPLGVGLA
    CWHEYNPLARALVQGEELHCAQFPGSLVGLPFEQTAVMRHLAEACFVIGSTG
    MLDEEDYALITPDASLHKAFQGGCHTAVISPEGRYLAGPLPDGGLAIELDLALITK
    RKMMDSDVGHRSPELLSLINRTAVAOEMASDPVSETSMATSMAGMSHVEE
    INHY"
    complement(3821..4303)
    /locus_tag="SYNM1426"
    /complement(3821..4303)
    /locus_tag="SYNM1426"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="conserved hypothetical protein"
    /protein_id="CAE07941.1"
    /db_xref="GI:33638935"
    /translation="MPVDRPANOQGDYLVDEKVPDVKAPGEKALVTHTVAE
    GSLVNLQASRLITKGFETSVLVGPVTLGVRGFPKLGDAFDPGLNPNARLQK
    FMDGGRKYACRFALQALYHSEKALMPGITPVNPLDVLIVMHRKEGFLIDITWL
    "
    4490..5800
    /locus_tag="SYNM1427"
    4490..5800
    /locus_tag="SYNM1427"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="Predicted flavoprotein involved in K+ transport
    (COG2072)"
    /protein_id="CAE07942.1"
    /db_xref="GI:33638936"
    /translation="MLSHADVAFAKSDHAAVVVIGAGQALSVAYOLQGRIRPVYLE
    KHRIGVAMDOQRMDSCFLVTPMOCRLPDPYDQNGPEGFMRAEIVAYLQRPAAVG
    GDVEGVAVORLTPKSGVRLITSEGEAEHVVATGVAHRRHPLAERLASVYLG
    LDARAYNPALPEGPVLVVGSGGSOIALDLHLSGRVTHLVGSAFSPRVRGRD
    VVPLDMRGYAMPISDHADPRSYRATNNVYLGQGGREIDLRHATGEMRLHGELA
    TITADHGFADDLAIGNLDQADAVYCRIRSSIDSVYIQGGIEALFEPPIECPQPSMA
    DPGLDSRDLPAAVICTGYRSDPSWIDAVFDGAGPAHERGVTSAGLYFGLPWL
    HTWSGRCFVGSDDADYIARLISLRQRDAQSORELCTAIIIGS"
    complement(5808..6734)
    /locus_tag="SYNM1428"
    complement(5808..6734)
    /locus_tag="SYNM1428"
    /codon_start=1
    /locus_tag="SYNM1428"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="hypothetical"
    /protein_id="CAE07943.1"
    /db_xref="GI:33638937"
    /translation="MPMRNSCARPFTSVYRTSGPTSTVULLQSGMNGSFE
    SASRSASTHEFDGLAEVNVSHCEIRIATVROLGRRALILEITGGGTPDLAH
    GPEQLHQLLEALLRALAMEPLKINILIGSDWSVEDQHRFFAVLDRIIDVCAVNL
    EHTSRSLANWQMPVLEHRRPMLRLTADSHWCVAERLMTDPLPVQAMARVHI
    HARVGAHGSVSHDFAPAEWTEALEARSCWQFLEFDEKAPATITBEFGPDGYNP
    LQPSAEPAVDVASLNTQMSWLEATLHNSMR"
    complement(7022..7729)
    /locus_tag="SYNM1429"
    complement(7022..7729)
    /locus_tag="SYNM1429"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
```

```
gene
CDS
    /product="DnaJ domain-containing protein"
    /protein_id="CAE07944.1"
    /db_xref="GI:33638938"
    /translation="MGFDPROWTSGPSGASTNCDERVTSNVEALLAENDALREILQ
    RMORELRLRPOQRPRPEPTMNLIRPVQVOMWGEELAKQKQMSFEDAGLEALIERLN
    RRGPPANLDOQLRNLRYLDGIGSDLLAAVNPATLOTAVALAALGVASLELDED
    PRVVAELHORLGRPRPSGDFRADALAVLGEPAETAEALIKARHRLVKLHFDMKG
    SAEAFRRVNEAYQQLVN"
    complement(7748..8437)
    /locus_tag="SYNM1430"
    complement(7748..8437)
    /locus_tag="SYNM1430"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="short-chain dehydrogenase/reductase (SDR)
    superfamily"
    /protein_id="CAE07945.1"
    /db_xref="GI:33638939"
    /translation="MGNGERELMETTYLVGTANRGIGLAXCOOLQARGDAVYACROIS
    PELESIGVLEAGIBLSQSIDELVRIDCPLDGVITLNGILHSGLEDLDEAIR
    ROFEYNALAPLLARALVPQMPQSKLALTSRNGSIDNSGGSYGRMSKALNTA
    GKSIALIDLPKRGIVAILHPGLVRTGMIRFNPSSGVPDSDAVQILARIDALTLETSS
    FMHANGVILPW"
    complement(8695..8922)
    /locus_tag="SYNM1431"
    complement(8695..8922)
    /locus_tag="SYNM1431"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="conserved hypothetical"
    /protein_id="CAE07946.1"
    /db_xref="GI:33638940"
    /translation="WPESTAVQQPILHLSRDLQVKNQAAITADEVMAAOELGYANS
    GSDLLLDGQHEAGIRVTRVDHPPEYRYY"
    8916..9203
    /locus_tag="SYNM1432"
    8916..9203
    /locus_tag="SYNM1432"
    /codon_start=1
    /evidence=not_experimental
    /transl_table=1
    /product="hypothetical"

Query Match      33.0%; Score 647.4; DB 1; Length 349213;
Best Local Similarity 61.8%; Pred. No. 9,9e-178;
Matches 1065; Conservative 0; Mismatches 647; Indels 10; Gaps 2;

    136 ACCGGGCGCTTATTCCTGATGATAGCTGAAGCCCATGGGGTCAACACATTTTGGC 195
    287244 AGCGGCGCTGCGGCGCTGATGATGCGCTGCGGCGCCACGGGGTGGACACGATCTCGGT 287185

    196 TATCCCGCGCGGCGCATTTTCCCATCATGATGAACGTACCGCTTGAAGCGGCGGG 255
    287184 TACCGAGGTGGGCGATCTCGCGATTTACACGCGCTGCACTGATGAGCGAAAGC 287125

    256 GAATTTGACATATTTTGGTGGCGGCATGACAGAGAGCTTCCATGGCGCGGATGGGTAT 315
    287124 TGGGTGAACAACATCTGCTGGCGGACAGAGAGCGGACCAATGCGGCTGATGCTTAC 287065

    316 GCCAAGGCCACAGATGAAGGTGGAGTTGTTGGGTACATCTGGACCAAGGGCCACATAC 375
    287064 GCCAAGGGCACCGGGAAGGTGGGTGCTGCTTGGCACTTGAAGTCCAGGGGCCAAT 287005

    376 TTGGTGAACCGGATGGCATGGCATGGCATTTGACTCGGTGCCCATGGTGGTATTACTGA 435
    287004 CTGGTGAACGGGATGGCCACAGCCAGATGATCTGTGCGCATGGTGGTATCACCGGA 286945

    436 GAGTGGCGCGTGGCATATTTGATGCGATGCTTCCAGAAATTGACATTTTGGCATC 495
    286944 CAGGTTCGCCCGCGCGCATGCGGACCGATGCTTTTCAGAAACGATATTTTCGTATC 286885
```

QY 496 ACCTTACCGATGTTAAGCACTCTATGTGTAAGTAGTCGGCGGATATGCTCGCAAT 555
 Db 286884 ACCCTCCCATCGTGAAGCATCTCTGGTGTGCGGACCTCTGTGACCTGTGGTTGCAATC 286825
 QY 556 GTTACTGAGGCTTTTCATCTTGTAGACACCGGTCTGTCCCGGCGGTTTATTCATATT 615
 Db 286824 GTGGCCAGGCGTTTGTATCGCCGACAGGTCGCGCCGCGCGCTGTGTATGCAATTC 286765
 QY 616 CCCAAGATGTGGGCTTTAGAGAAATGTAGTACATTCCTCGACCCCGGTAGCTTAT 675
 Db 286764 CCCAAGATGTGGGCGAGAAAGTTCACTACGTGCGGTGGAACCGGTTCCTCATC 286705
 QY 676 CTACGGGTTATCGCCCAAGGTTAAAGTTATCCCGCAAAATTAATGCGGATTGCA 735
 Db 286704 CTGGCGGTTTCCATTCAGCCCAACCTCGCTCATGTGCGGTGACGCGCACTGGAT 286645
 QY 736 TTGTTGAGCAGCAGCCAGAAATCCCTTGTCTATAGTGGGGGAGGGGCGATGCGCGCAAT 795
 Db 286644 CTGATCGAGCAGGCGCAAGACCGCTGTCTGTACGTGGGTGCGGGCGATCTTGTCTGT 286585
 QY 796 GCCCATGCCAGGTGCAAGAAATTTGCGAAGGTTCCAGTTGCCGTAAACAACCCCTG 855
 Db 286584 GCCCAGCAGAGCTGCGAGTGTGCGCGGAGCTTACAGCTCCGATCACCAACCTTG 286525
 QY 856 ATGGGAATGGGGCTTTTACGAAAACCATCCCTTGGGTGGTATGTTGGTATGAT 915
 Db 286524 ATGGGAATGGTCCCTTTATGAAACGATGCTCTCTGGTGGGATCTCGTATGAT 286465
 QY 916 GGCCACCGCTATGCCAATTTGCGGTACGAGATGTAATTTGTTGTTGCAAGTGGGCGC 975
 Db 286464 GGCCACCGCTATGCCAATTTGCGGTACGAGATGTAATTTGTTGTTGCAAGTGGGCGC 286405
 QY 976 CGTTTCAGACGACGGGTAATCTGCAACTAGACGAAATTTGCTAGCCGCGCAAACTATT 1035
 Db 286404 CGTTTCAGACGACGGGTAATCTGCAACTAGACGAAATTTGCTAGCCGCGCAAACTATT 286345
 QY 1036 CACATTTGACATGACCCCGGAGGTGGGAAAAACAGGGCTCCCGATGTGCCATTGTG 1095
 Db 286344 CACTTCGAGATGATCCGCGGAGATCGTAAACCCAGAGCTGATGTTGCCGTCTC 286285
 QY 1096 GGGAGTGTACGCCCATGTTTGAACAAGCTTTTGACGCGGGCCCGGAAATGATTAACCC 1155
 Db 286284 GGTGTCTTCGGCTTGAAGCTGTGCGGAGTGTGAGATCAGCTTGCA-----CGC 286234
 QY 1156 ACCCATCCCATPACCAACCAAGCATGTTTAAATGCAATGATTCAGTGGCGGACCGATAC 1215
 Db 286233 ACGGCGGAACCCCGCATGTCAGCTGTGAGGCCCATCAACCTGGAAGATGTTAT 286174
 QY 1216 CCGCTCCAGGTGCCCACTATGAGATCTATTGCCCCCAAGAGTAGTACAGCAATT 1275
 Db 286173 CCGCTCCAGATTCGCCCAAGGAGGAGATCTACCCCAAGAGTAGTCTGTGCGGTG 286114
 QY 1276 GGTGCCGAGGCGCCCGATGCTACTACACACCGATGTGGACAAACCAATGTGGCG 1335
 Db 286113 AAGGATTCGCGCCCGCAATGTCACACGAGCTGTGTGACGATCAGATGTGGCT 286054
 QY 1336 GCCCAGTTTGAACATGCGCCCCCGCGAGATGTTCCAGTGTGCTGTGGTACGATG 1395
 Db 286053 GCGCAGCAGCTGCGGACGAGGCGGCGGATGATGAGCGCTGTCTCGGACGATG 285994
 QY 1396 GCGTTTGGTTTACCTGCGCCCATGAGAGCCAAAGTGGGAGTGGGGGA CGAGCGTCAATT 1454
 Db 285993 GCGTTTGGAGATGCGCGCTGCAATGGGGCTCAGGTGCGCATGCGGATGCGCAGGTGTG 285934
 QY 1455 TGCATCAGTGAATGATGAGCTTCAATGAACTTTTACAGAACTGGGAACTCTGAGCCAG 1514
 Db 285933 TGCATTCGCGGTGATGCGCAGATCTCTATGAATCCAGGAATCTCGAACTCTGCGCC 285874
 QY 1515 TACGACATTCAGGTTAAATCTATTATCTCATTAACGTTGTGGCAGGGGATGTTGCTGAG 1574
 Db 285873 TATGCGCTGCGGTGAGGTGTGATGTCACCAACCACTGCAAGGCGATGTGCTGAG 285814
 QY 1575 TGGCAACAACTTTCTACGAAGAAGTATTCTGTTTAAATGATGCCAGGCGATGCA 1634

Db 285813 TGGCAGGAAAGCTTCTACGACGAGCGCTATTCCGCCCTCGACANTGCTCATGCAATGCC 285754
 QY 1635 GACATTAATCTCTCTCTGTGAAGCTTATGCACTAAGGATATTACTGTGCCAAAGCGGAA 1694
 Db 285753 GATTTCATCGCCCTGAGCCCGCTTCTTGGGGTGTGATGCGTCAAGATCACCGACCTGAG 285694
 QY 1695 GATTGGCCCCCGGCGATGCGCAAAATGCTAGCCCAACAATGTCCTGTGATGATGATG 1754
 Db 285693 TTGCTGACACCGCACTCTGTCAGCCCTGCAATCACCAACCCGACGATGATGATG 285634
 QY 1755 GTGTCAAAAAAGATGAAAATCTTTACCTTATGATTTGCCCCGCGATGAGTAAATGCCAA 1814
 Db 285633 CACGTGCGGGGAGAGAACTGTAACCCGATGATGTCCCGCGGCAAGACAAATGACAG 285574
 QY 1815 ATGCTAGTTTACCGGAAGTCCCGGTACGACGACAAATGCTCC 1856
 Db 285573 ATGATGGATCTCCCTTCGATCCGAGCTGGCAATGGGACCC 285532

RESULT 8
 PSP224709
 LOCUS
 DEFINITION Porphyridium sp. chloroplast gene encoding acetylhydroxy acid synthase.
 ACCESSION AJ224709.1 GI:3157427
 VERSION AJ224709.1
 KEYWORDS acetylhydroxy acid synthase, ahas gene, chloroplast Porphyridium sp.
 SOURCE Porphyridium sp.
 ORGANISM Eukaryote; Rhodophyta, Bangiophyceae; Porphyridiales; Porphyridiaceae; Porphyridium.

REFERENCE
 1 Lapidot, M., Ravesh, D., Sivan, A., Arad, S., and Shapira, M.
 The chloroplast-encoded Ahas gene of Porphyridium sp. is distinct from the nuclear-encoded gene of green algae and plants

JOURNAL
 Unpublished
 2 (bases 1 to 2338)

REFERENCE
 2 Direct Submission
 Submitted (23-FEB-1998) Ravesh D., Life Sciences, Ben Gurion University of the Negev, P.O.Box 653 Beersheba, 84105, ISRAEL

FEATURES
 source
 1..2338
 /organism="Porphyridium sp."
 /organella="plastid:chloroplast"
 /mol_type="genomic DNA"
 /db_xref="taxon:77580"
 /note="rhodophytes"
 544..2307
 /gene="ahas"
 544..2307
 /gene="ahas"
 /codon_start=1
 /transl_table=11
 /product="acetylhydroxy acid synthase"
 /protein_id="CAA12081.1"
 /db_xref="GI:3157428"
 /db_xref="GCA:063181"
 /db_xref="SPRMBL:063181"
 /translation="MTHIRKSNVQEOGTGAFALIDSLVRHKVHIFGPGATILPIYDE
 LYKMEQGYIKHILVHEOGAAMADGVARATGEGVCPATSGPAMTLVGTATAM
 KSIPYITGVGVRSPIGTDAFOEVDIFGILTVGHSYVIRDPRIPIVAVPSIA
 KGRPGPVLIDVKDGLTPEYOVYVPEEARIPGRDVLVAPSSQIHSIQLDEAN
 OPTLVYGGATISGADHLTYKLVNOKRPIITTTMGKTIIDPQNPLALGMLMGHAY
 ANFAYECSHLLITLGRFDRVTGKIDEPACAKYIHDVIDRAVGKNIPOVALVGD
 ISLVBCWLLVLDRLQLDSDSHRKHSEIFRROBYPLIVPLVQTLSPQSIANIS
 QIMPDAVFSTDVGOHQMAAQFVKTLPRMILSSSGIGTGYGLPAIAGIKAIYPSPV
 VCTITGDSFQNMNIOELGTIAQYKLIKIIINNKWQGMVWQCAFYAGKARYSHMED
 GAFNFAVLAKSFGIDGOSISTROEMDSLNTIITKXGPMVIDCKVIEDNCVPMVAPG
 KSNAMQIMIGDKSNNEIKIKE"

ORIGIN

Query Match 32.6%; Score 638.4; DB 8; Length 2338;
 Best Local Similarity 61.8%; Pred. No. 2e-175;
 Matches 1050; Conservative 0; Mismatches 646; Indels 4; Gaps 2;

127 CAAGCGCAACCGGCGCTTATATCCGATGAGTCCCTGAAACCGCTGGGGCTCAACAC 186
 571 CAAGCAACACCGGCGCTTATATCCGATGAGTCCCTGAAACCGCTGGGGCTCAACAC 630
 187 ATTTTGGCTATCCCGCGGCGCAATTTTGGCTATGATGATGATGATGATGATGATGAT 246
 631 ATTTTGGCTATCCCGCGGCGCAATTTTGGCTATGATGATGATGATGATGATGATGAT 630
 247 GCGCGCGGCGCAATTTGAGCATATTTTGGTGGCCATGACAAAGAGCTTCCCATGCGCG 306
 691 GAGCAAGGATATATTAAGCATATATGCTTGAACAACAAGGCTGTGACATGCTGCA 750
 307 GATGGATATGCGACAGCAGCAGTAAAGTGGAGTTTGTTCGCTACATGACGACGAG 366
 751 GATGGATATGCGACAGCAGCAGTAAAGTGGAGTTTGTTCGCTACATGACGACGAG 810
 367 GCGATTAATCTGGTACCGGCGATTTGCCATGCTCCATTTGACTGCTGGCCATGCTGTG 426
 811 GCGCTATATTAAGCAAGGATTTGCAACAGCTCATGATGATGATGATGATGATGATGAT 870
 427 ATTTGAGAGAGTGGCGCGCTGCGCATGATGATGATGATGATGATGATGATGATGATGAT 486
 871 ATTTGAGAGAGTGGCGCGCTGCGCATGATGATGATGATGATGATGATGATGATGATGAT 930
 487 TTTGGCATCACCTTACCGCATGCTTAAGCATCTCTATGCTGATGCTGCTGCTGCTGCTG 546
 931 TTTGGCATCACCTTACCGCATGCTTAAGCATCTCTATGCTGATGCTGCTGCTGCTGCTG 990
 547 GCTGCATTTTATCTGAGGCTTTCATCTGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTG 606
 991 CTTGAATATGCTGAGGCTTTCATCTGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
 607 ATTCATTTTCCCAAGAGTGGGCTTAAGAGATGATGATGATGATGATGATGATGATGATGAT 666
 1051 ATTCATTTTCCCAAGAGTGGGCTTAAGAGATGATGATGATGATGATGATGATGATGATGAT 1110
 667 GACCTTATCTACCGGCTTATGCGCCACGCTTAAGGATATCCCGCAAAATTAATGCG 726
 1111 GAAAGCAAAATTAAGGATTTTCCGATGATGATGATGATGATGATGATGATGATGATGAT 1170
 727 GCATTCATTTGAGAGGCTTTCATCTGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
 1171 TCAATTCATTTGAGAGGCTTTCATCTGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
 787 GCGGCAATTCGCGCATGCTGCGAGTGAAGATTTGCGGAAGCTTCCGCTGCGGCTACA 846
 1231 ACTTCGAGGCTCATGACTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1290
 847 ACCACCGCTGAGGAGATTTGGGCTTTCAGCAAAACATCCCTTTCGCTGCGGATGCTG 906
 1291 ACTACATTAATGAGGATTAAGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 1350
 907 GGTATGATGCGCACCGCTATGCAACATTTTCCCTGACGCGAGTGTGATTTGTTGATGCA 966
 1351 GGCATGCAATGCTACTGCTTATGCAAAATTTTCCCTGACGCGAGTGTGATTTATTAATACC 1410
 967 GTGGGCGCGCGCTTTCAGACCGGCTTTCAGCAAAACATCCCTTTCGCTGCGGAGCGCC 1026
 1411 TTGAGAGCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
 1027 AAAGTAATTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
 1471 AAAGTAATTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530
 1087 CCCATTTGGGGGAGTACGCGCATGTTTAAACAAGCTTTTGACGCGGCGCGGAGATTG 1146
 1531 GCCATTTGGGGGAGTACGCGCATGTTTAAACAAGCTTTTGACGCGGCGCGGAGATTG 1587
 1147 GATTACCCACCATCCCATACACACGAGCATGTTAATTCGATGATGATGATGATGATGATGAT 1206

1588 AACTTGCAATTAAGATAGACCATGCTTGGTTCATGCGATGAGCGGTATATTTAGTGGGCG 1647
 1207 ACCGATTAACCCCTCCAGAGCGCCCATATGATGATGATGATGATGATGATGATGATGATGAT 1266
 1648 CAAGATATCCATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1707
 1267 CAAGAAATTTGGTGGCGAGCGCCGAGTCTTACACACCGATGATGATGATGATGATGATGAT 1326
 1708 GCGAATATTTAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1767
 1327 AATGAGGCGCGCGCTTTCAGACAGCGCCCGCATGATGATGATGATGATGATGATGATGATGAT 1386
 1768 AATGAGGCGCGCGCTTTCAGACAGCGCCCGCATGATGATGATGATGATGATGATGATGATGAT 1827
 1387 GGTACGATGAGCGCTTTCAGACAGCGCCCGCATGATGATGATGATGATGATGATGATGATGAT 1445
 1828 GGCACAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1887
 1446 GGGCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
 1888 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1947
 1506 CTAGCCAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
 1948 ATAGCCCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2007
 1566 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1625
 2008 GTAGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2067
 1626 GGCATGCGACAGATTAATTCCTCTGTAAGCGCTTACGATGATGATGATGATGATGATGATGAT 1685
 2068 GGAGCACCTTAATTCCTCTGTAAGCGCTTACGATGATGATGATGATGATGATGATGATGAT 2127
 1686 AAGCGGAAATTTGGCGCGCGCGCATGCGCGCAATGCTAGCCACATGATGATGATGATGAT 1745
 2128 ACTAGGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2187
 1746 ATGAT 1805
 2188 ATAGACTGTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2247
 1806 AATGCCCAATGCTAGGTTT 1825
 2248 AAGCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2267

RESULT 9
 BX572098/c 346683 bp DNA circular BCT 23-SEP-2003
 LOCUS
 DEFINITION
 Prochlorococcus marinus MIT9313 complete genome; segment 4/7.
 ACCESSION
 BX572098.1 GI:33640689
 VERSION
 complete genome.
 KEYWORDS
 Prochlorococcus marinus str. MIT 9313
 Prochlorococcus marinus str. MIT 9313
 Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
 ORGANISM
 Prochlorococcus.
 1 (bases 1 to 346683)
 Roca, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P.,
 Ahlgren, N. A., Arellano, A., Coleman, M., Hauser, L., Hess, W. R.,
 Johnson, Z. I., Land, M., Lindell, D., Post, A. F., Regalado, M., Shah, M.,
 Shaw, S. L., Steglich, C., Sullivan, M. B., Ting, C. S., Tolonen, A.,
 Webb, E. A., Zinser, E. R. and Chisholm, S. W.
 Genome divergence in two Prochlorococcus ecotypes reflects oceanic
 niche differentiation
 Nature 424 (6952), 1042-1047 (2003)
 JOURNAL
 MEDLINE
 22825698
 12917642
 PUBLISHED
 2 (bases 1 to 346683)
 REFERENCES
 Larimer, F. and Roca, G.
 Prochlorococcus genome consortium

Db	278847	CCGACGAGATTTTCCACGATTCGCCGCTCCAGAACTTTCTCTCAATTGAGCGGCTCTGAT	278788b
QY	736	TTGTTGAGACAGCCAGAAATCCCTTGCTTACGTAGGGGAGGGGCGATTCGCCCAAT	795
Db	278787	TTATATGATGATGCTCAGAGGCTCTCTCTTAACGTGCGGAGAGATTATTTCTGCCGGG	278728b
QY	796	GCCATATGCCAGTGTGAGAAATTTGCGGAAAGGTTCCAGTTGCCGATTAACAACCCCTG	855
Db	278727	GCTCAGAGAGCTTAAAGGCTTTTCAGAAAGTCCACGATACCTGTCAACAACAATTCG	278668b
QY	856	ATGGGAATTTGGGGCTTTTACAGAAAAACATCCCTTCGGTGGGTATGTTGGGATATGAT	915
Db	278667	ATGGGTAAGGGGGCGTTTGTATAGGCAACATCTTTGGCTTTGGGAAATGTTGGAAATGAT	278608b
QY	916	GGCACCGCTATGSCCAATTTGCCCTTCACGATGTGATTTGTTGATTTGACGTGGGGCC	975
Db	278607	GGCACTGCTATGCCAACTTTTGCAATTTCTGAAATGTATTTGTTGATCGCTGTGGTCA	278548b
QY	976	CGTTTGAAGACCGGGTAACTGGCAAATAAGCAATTTGCTAACCGCGCCAAAGTAAT	1035
Db	278547	CGTTTGAATGACCGTGTTCACAGCAAGCTTGATCTTTTGGCCCCCGCAAGATGAT	278488b
QY	1036	CACATTGACATCGACCCGCGGAGGTGGGAAAAACAAGGCTCCCGATGTGCCCATTTGTG	1095
Db	278487	CATTTTGAATTGATTCAGACAGAAATGATAGTAAGAACGACGCTCGATGTTGGCGTATTCG	278428b
QY	1096	GGGAGATGACGCCATTTTGTAGACAAGCTTTTGCACGGCGCGCGGAATTGATTAACCC	1155
Db	278427	GGTGAATGTTGGTGTGAGCTTTGTGTGAAGTTATTCGACTTGAGTAAGCAAAATCTGT---	278377b
QY	1156	ACCATGCCCATATCCACCCAGGCATGTTAAATGSCATGATTCATTGGCGACGATTAAC	1215
Db	278371	-----TGAGTTGAGACTTCACGCTGGCTACTACCTGATTTGAGAGTTGGAAACCTTTAC	278312b
QY	1216	CCCCCTCAGGTGCCCCACTATAGAAATCTATTGSCCCCCCGAGAGTGTACACGAAATT	1275
Db	278316	CCCTTATGATGATCCGCTGAGGAGGGGGCCAAATTTATCTTCAGAGGTGCTCTTGGCGATA	278257b
QY	1276	GGTGGCAGGGCCCCCGATGCTACTACACACACGATGTGGACAACCAATGTGGGCG	1335
Db	278256	AGGGACTTGCTCTTGAAGCTCTATGAGCTTACATACACTGATGTGGGCGACATCAATGTGGCG	278197b
QY	1336	GCCCAAGTTTGTGAACAAATGGCCCCCGCGAGATTTCCAGTGTGCTTGGGTACATG	1395
Db	278196	GCTCAATATTTACGAAATGGAGGCTGTGTCAGTGAATCACTAGACGGGGATTGGGAACATG	278137b
QY	1396	GGCTTTGTTTACCTGCGGCGCATGGGAGCAAAAGTGGGAGACGAGC--GGTCAAT	1454
Db	278136	GGTTTGGAGTGCACAGCTCTATGTGTTCAGTTGCTTTGCCAGATGAACAGTTGTT	278077b
QY	1455	TGCATCAGTGGAGATCCAGCTTCCAAATGAATCTTCAAGAACTGGGAACCTTACCCAG	1514
Db	278076	TGTATCGCTGTGATTTCCAGCATCTCGATGAATATTTCAGAGTTGGGAACACTGATTTGAG	278017b
QY	1515	TAGCACTATCCAGGTTHAACTATATNTCCATTAACGTTGGCAGGGGATGATGTCGACG	1574
Db	278016	TACAACTTCCACACAGATGTGATTTGAACAACACTGGCAAGGATGTTCCTCAG	277957b
QY	1575	TGGCAACAACTTTTACAGAAAGCTTATTTGCTTTAAACATGTCCACAGGGCATGCCA	1634
Db	277956	TGGAGCAAAAGCTTTTACGAAAGTGTATTTCAGCACATGACATCTTCGGGACATGCT	277897b
QY	1635	GACATTATTTCTCTGTGTAGAGCTTATGATCAACAGGTATTTACGTGCCGAACGGGAA	1694
Db	277896	GATTTCTGTGCATTTGCGAAGGCTTTTAAAGTTGCGGGGAGTTGATATCTGAGGGAAG	277837b
QY	1695	GATTTGGCCCCGCGCATTCGCCAAATGTCTAGCCCAATGTCTGTGATGATGATGTG	1754
Db	277836	GATTTCAAGTCAAGTTTGAAGACAGGCGCTTGGCATCCAGGTCGATCTCATGACGTTG	277777b
QY	1755	GTGGTCAAAAAAGATGAATCTTTACCTTATGATTTGCCCGGATATGTAATGACCAA	1814

[illegible]

	CVSQSDSEFCMMMDLGTIAAYKPIKIKVIIINNRMOGVVRVMOQAIFYERISHSMTEG
	AHPNQSLEAFEGICATFVNKRNMPESSLKDMAKPKPGVLILDCVTENENCYPRVAPEK
	SNAIMIGIAPQRRETAISNVSRNI "
gene	2485 . .3336
	/gene="argB"
CDS	2485 . .3336
	/gene="argB"
	/EC.number="2_7_2.8"
	/note="putative"
	/citation=[1]
	/codon_start=1
	/product="N-acetylglutamate kinase"
	/protein_id="AAO3053.1"
	/db_xref="GI:343131"
	/translation="MLINTERVKVLSDVTLLOKPSSRIIVIKYGAAMKNOKLDHVLI
	SLDVFYSFIQLRPILVHNGGPRLINFMIDDKITLPKPNGRVDOPRMDIVEMLVNER
	VNKDLVASINKKGSGSVGLSGKDGLLTSPSKPNTGFGEVONUTNLLETIIANN
	YIPAVASAADKKQSQSINIADTVAGETARLNAEKIILTDPGITLRNSSDATLTIS
	HLSIQEARLDLTKAIVISGMIPRVNCIRSLAQGVASHAILEDRIDHALLETILDQG
	IGSNLVV"
gene	3504 . .3577
tRNA	/gene="trnfm(CAU)"
	3504 . .3577
	/gene="trnfm(CAU)"
	/product="tRNA-Met"
	/note="putative"
	/citation=[1]
	/anticodon=(pos:3538 . .3540,aa:Met)
gene	3583 . .3654
tRNA	/gene="trna(GGC)"
	3583 . .3654
	/gene="trna(GGC)"
	/product="tRNA-Ala"
	/note="putative"
	/citation=[1]
	/anticodon=(pos:3616 . .3618,aa:Ala)
gene	3699 . .3788
tRNA	/gene="trns(GCU)"
	3699 . .3788
	/gene="trns(GCU)"
	/product="tRNA-Ser"
	/note="putative"
	/citation=[1]
	/anticodon=(pos:3734 . .3736,aa:Ser)
gene	3928 . .3992
tRNA	/gene="trnd(GUC)"
	3928 . .3992
	/gene="trnd(GUC)"
	/product="tRNA-Asp"
	/note="putative"
	/citation=[1]
	/anticodon=(pos:3962 . .3964,aa:Asp)
ORIGIN	
Query Match	30.8%; Score 603.6; DB 8; Length 4099;
Beech Local Similarity	60.5%; Pred. No. 3.Se-165;
Matches 1029; Conservative	0; Mismatches 669; Indels 4; Gaps 2;
Oy	132 GCAAACCGGGGCTTATTCCTGATGATGATTACCTGAAAAGCCATGGGCTCAACACATTTT 191
Db	2018 GAATACTGTCGTTTGCTCCTTTTAGATAGCATTTGTCGTAAGAGATTATTCATATATT 1959
Oy	192 TGGCTATCCCCGGCGGGGCAATTTGCCCATCTATGATGAAGAAGCTTGACCCTTTGAAGCGGC 251
Db	1958 TGATTACCTCGTGCGGGCTATTTTTACAATTMTGATGAACCTTATGCTTGGGAGAACT 18599
Oy	252 GGAGGAATTGAGCATATTTTGGTGCCCATGAACAAGAGACTTCCATGCGGCGGATGG 311
Db	1898 TTGATTATTCAGAAATATTTCTTGTAACACATGAACAAGAGTGCTTCAATGCTGAGATGC 1839
Oy	312 GTATGCCAGAGCCACAGTAAAGTGGAGTTTGTTCGTATCATCTGGACCAGGGGCGAC 372
Db	

QY 372 TAACTGTGATGACCGGCAATGGCCATTTGCCATTTGGATCGGTGGCCATGGTGTGATTTAC 431
 Db 1778 TAAATTTAGTATCCCGCAATGACTGCGCAATTTGACTCGGTCTTATTTACTATTTAC 1719
 QY 432 TGGAGAGTGGGCGGTCATGATTTGGTAGGATGCTTTCCAGAAATTTGACATTTTGG 491
 Db 1718 TGGCAAGTGGGAGGACCTTTTATAGGAACAGATGCAATTCAGAAAGTTGATATTTTCG 1659
 QY 492 CATACCTTACCCATGCTTAAAGCACTCTTATGTGTACGTATGTGCGCGCATTTGCTCG 551
 Db 1658 CATTAACATTAATTTGTAAGCAATTCATATGTATGTCGGTACCTTAAAGAACATGCTCTAG 1599
 QY 552 CATGTTACTAGAGCTTTCCATCTTGTAGACACCGGTCGTCCGGCGGCTTTTGAATCGA 611
 Db 1598 AATTTGTTGCTGACGCTTTTACATTTGGCAACATGTATGACCTGCGCCATTTTATTTGA 1539
 QY 612 TATTCCCAAGATGTGGGCTTAAAGAAATGTAGTACATTCCTCTGACCCCGGTGACGT 671
 Db 1538 TGTACCAAAAGATGTGGACTAGAAAATTTATTTCTCTGTTGAGCAGAGGCAAGT 1479
 QY 672 TAACTACCGGGTATGCGCCCAAGTTAAAGGTAATCCCGCAAAATTAATGCGGCAT 731
 Db 1478 TAAATCCCTGTGCTGACAGCTCTATCTTAATCTTAAATCCCTGCAAAATTAATGCTG 1419
 QY 732 GCAATTTGTGAGAGGCGCAAAATCCCTGTCTAGTAAAGGAGGAGGCGATGCGCCG 791
 Db 1418 AAAATGATTCACAAATCAAGTCAAGCATTTATTAATTTGTTGAGTGTGAGCTATTAATTT 1359
 QY 792 CAATGCCATGCCAGGTGACGAAATTTGGGAAAGTTCAGTTGCCGCTAACACAC 851
 Db 1358 TGAATGCTACATTAATTAAGATTAATTTGACTTATTAATTAATTAATTAATTAATTAAT 1299
 QY 852 CCTGATGGGAATTTGGGCTTTTGAAGAAATCCCTCTGCTGAGTGTGTTGTTGTTGTTGTT 911
 Db 1298 TCTTATGGGCAAGATTTCTTAAATTTAGAGATGATGATTTTGTCTAGTATGCTCGGAAT 1239
 QY 912 GCATGGGCAAGCTTATGCAATTTGCGGTCAGGAAATGTGATTTGTTGATTTGCAATGGG 971
 Db 1238 GCATGGGCAAGCTTATGCAATTTGCGGTCAGGAAATGTGATTTGTTGATTTGCAATGGG 1179
 QY 972 GGCCTGTTGCAAGACCGGCTTATGCAATTTGCTAGGCGGCGCAAGT 1031
 Db 1178 AGCTAGATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119
 QY 1032 AATTAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
 Db 1118 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1059
 QY 1092 TGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
 Db 1058 TGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1002
 QY 1152 CCCCACCCATCCCATACCAACCCAGGATGTTAAATGCAATGATGATGATGATGATGATGATGAT 1211
 Db 1001 TAAACCTTATTCAGAGCAAAATTAATCTTGGCAAGAAATGATGATGATGATGATGATGAT 942
 QY 1212 TTACCCCTCCAGGTGCGCCATTAAGATTAATGATGATGATGATGATGATGATGATGATGAT 1271
 Db 941 AATACCTCTTTTATTAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 882
 QY 1272 AATGATGCGAGGCGCCCATGCTTACTACACACCGATGATGATGATGATGATGATGATGATGAT 1331
 Db 881 AAAAAATCAATTAACGAGATGCTTATTTTACTAGATGATGATGATGATGATGATGATGATGAT 822
 QY 1332 GCGCGCCCATGTTTGAACAATGAGCCCGCGCATGATTTCCAGTGTGCTGGTGTAC 1391
 Db 821 GTACGCTCAATTTCTTAAAGTGAATTAAGCAATTTGATTTCTAGTGTGCTGATGATGATGAT 762
 QY 1392 GATGGCTTTGTTTACCTGCGCATGAGCAAAATGAGTGGAGTGGGAGCGAGG-GGT 1450
 Db 761 AATGGCTATGTTTGCACAGCTATTTGTTGCCAAGTACACACCTTAACCAATTAAT 702

QY 1451 CATTGCAATGAGAGATGCGAGCTTCCAAATGATTTGAGGAACCTGGACCTAGC 1510
 Db 701 AATTTGTTAAGCGGTGATTAAGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 642
 QY 1511 CCAGTACGACATCCAGTTAAACTATTAATTTCAATTAACGTTGGAGGAGATGATGCG 1570
 Db 641 CCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 582
 QY 1571 TCAGTGGCAACAATTTCTTCAAGAAAGCTTATTTGCTCTTCAATTAATGATGATGATGAT 1630
 Db 581 ACAGTGGCAACAATTTCTTCAAGAAAGCTTATTTCAATTAATTAATTAATTAATTAATTAAT 522
 QY 1631 GCCAGACATTAATTTCTTCTGTGAGAACCTATGAGCATCAAGGATTTAATGATGATGATGAT 1690
 Db 521 TCTTAATTTTCAAAAGCTGCTGAGAGCTTTGATTAATTAATTAATTAATTAATTAATTAAT 462
 QY 1691 GGAATTTTGGGCGCGGATGCGCGGAATCTGACCCCAATGATGATGATGATGATGATGATGAT 1750
 Db 461 ACAAATTAATTAATTTCTTAAAGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 402
 QY 1751 TGTGTGTGCTCAAAAAGATGAAATCTGTTACCTATGATTTGCGCCGCAATGATTAATG 1810
 Db 401 TTGCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 342
 QY 1811 CCAATGCTAGGTTTACCGGAA 1832
 Db 341 ACAAATGATAGGAATAGCAAAA 320

RESULT 11
 PUB38804
 LOCUS PUB38804 191028 bp DNA circular P1N 27-MAR-1998
 DEFINITION Porphyra purpurea chloroplast, complete genome.
 ACCESSION U38804
 VERSION U38804.1 GI:127652
 KEYWORDS chloroplast Porphyra purpurea
 SOURCE Porphyra purpurea
 ORGANISM Porphyra purpurea, Bangiophyceae; Bangiales; Bangiaceae; Eukaryota; Rhodophyta;

REFERENCE 1 (bases 1 to 191028)
 AUTHORS Reith M.E. and Munholland J.
 TITLE Complete nucleotide sequence of the Porphyra purpurea chloroplast genome
 JOURNAL Plant Mol. Biol. Rep. 13 (4), 333-335 (1995)
 REFERENCE 2 (bases 1 to 191028)
 AUTHORS Reith M.E.
 TITLE Direct Submission
 JOURNAL Submitted (17-OCT-1995) Michael E. Reith, Marine Biology Section, NRC Institute for Marine Biosciences, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada

FEATURES
 source
 1..191028
 /organism="Porphyra purpurea"
 /organelle="plastid:chloroplast"
 /mol_type="genomic DNA"
 /strain="Avonport"
 /db_xref="taxon:2787"
 181..1053
 /gene="chl"

CDS
 181..1053
 /gene="chl"
 /codon_start=1
 /product="prochlorophyllide reductase iron-sulfur ATP-binding protein"
 /protein_id="AAC08073.1"
 /db_xref="GI:127653"
 /translation="MKLAAYGKGGIGKSTTSCNISVALSKRKKVQLIGDPRKHSFT
 TLNGLIPITIDLOADYHEDVWPEDVYKGYGVDCVBAAGSPGAGCGGVVE
 TVKLKELNAEDVDILPVLGAVVCGFAAPANTYDCLITDNGFDALFANRFA
 ASVERKARTSLRAGLVGRTRERDIDKVIDCPVPEVLEVLIEDIVSRVKGKT
 LFEVASEDTNLAAYCDYLYNADQLIRPBGVFKEAADNELSLSDPLNPRSNNT
 KSTADEEELDMV"

gene	1095. .2402 /gene="ch1n" 1095. .2402 /gene="ch1n" /codon_start=1 /product="proteolophyllide reductase ch1n chain" /protein_id="AAC08074.1" /db_xref="GI:1276654" /translation="MSIVSDALTFECETGNYHTFCPISCVSWLYOKIEDSPFLVIGT KTCGYFLGNAMGVMIPAPRYAMAELEGGISAKLNDYGEIRLCLQIKDRNSVIF WIGTCTTEIITKMDLEGIAPKLBARIPIYVARANGDIATQGEDVTLAMORCS HHYIENINPPLVPLFGSLDPVVTQULWELAKQGIIVSGMLPESKRYTELPIQEGY VAGVNPFLSRATTLMRKRKTLIGAPPIPDQTRAMIEKICVFSIKEGLAEREK DIWDSLEDYIALIRGKSVFEMGDMLLEVSILARPLRCGMATVEIGIPIYMDRYOALR ALLKSTCDMMVMVMPPIVEKPDYNQDLRIDLKPDIYITGMALANPLERAGINTKMS VEFTFAOIHGFTNARDILELVTPLRRLSLSELGWDIYSKQG" complement (2454. .3929) /gene="ORF491" complement (2454. .3929) /gene="ORF491" complement (2454. .3929) /gene="ORF491" /codon_start=1 /protein_id="AAC08075.1" /db_xref="GI:1276655" /translation="MNFPIQDRLILKSRYPILINTREERLEYIKNOFCIDNKKV YCPDPDYTSNPNDNGYAKRNPILALDFPIKLTPTPLNLTIKDPSFNDGLIK LRNSQIKTQSKNIIIVSCRNIPLYLSDVITTVDPPLPMLSIKKEITRLSIALGI VLDELVNMLAKSQGSLVERIKRTITLQYQDLSRLPIIIEKRGIDNQTLL EFYPKKATEDIIGLDALSKWLKRSRFSKQSNFYGIPTPKGLLVIGIQGSLTLA KAIANDWTLPLRLDIGLFGGLVSESRKREWVTLAEGSPCVLWIDEIDKFSNL YSQGSQSGTAVRPFETIIVSEKTPPVAVANTIONLSEMLRKGRPEIFLIDP NMEREELFOJHLAIRPKSNQNNITQSLTCKRFSAGEIQAIVESMTABEENE FSTEDIKIALBOFIPLATDKQVESIQAMANGRRARASL" complement (4047. .4187) /gene="rpl134" complement (4047. .4187) /gene="rpl134" /codon_start=1 /product="50S ribosomal protein L134" /protein_id="AAC08076.1" /db_xref="GI:1276656" /translation="MTKGTIQGSKRKIRISGFAPRMKTSGRSILNERRRGRKKIM AS" 4333. .4854 /gene="yycf37" 4333. .4854 /gene="yycf37" /note="hypothetical chloroplast ORF 37." /codon_start=1 /protein_id="AAC08077.1" /db_xref="GI:1276657" /translation="MTVLTPIFLYSILTLPLTFESWLSIQOLKTPILLESQOYFLDK SKNDELIGEDPAPRSKVTIAKKYPTRAITESHVLTKKNSLESPEKNAIITAKTNMG FTIYADQKRLAKNYERPAIEVDGNYVALNNLAKIYEDTKNIIKAEALYDKVNLTK SNKILNTHKRFIK" complement (4889. .4962) /gene="trnp" complement (4889. .4962) /gene="trnp" /product="tRNA-Pro" complement (5020. .5544) /gene="ORF174" complement (5020. .5544) /gene="ORF174" /codon_start=1 /protein_id="AAC08078.1" /db_xref="GI:1276658" /translation="MTFWKTFPQNSYVGESENNPKFLIDJKTFDKIELKDYVLNST KNMILWVVLHPDFOGLGKVVWHLIKQLROALEISTITLFAEPDIVISPYRLGSI KDPGVKGMFWYPR" 5807. .6364 /gene="psaf" 5807. .6364 /gene="psaf"
gene	6377. .6505 /gene="psaH" /codon_start=1 /transl_except="pos:5807. .5809, aa:Met) /product="Photosystem I reaction centre subunit IIT" /protein_id="AAC08079.1" /db_xref="GI:1276659" /translation="MKKATFITCLAVLLVSNPIVNAEVALIPCKDSAPFNKRMVN SVKLOARLAKYDANPTPALANKOIEKTKTFATYGRAGLTCGDDGLPHLISDGRMS RAGDFEYFGLPLFLYITGMIGWVRGLLSVAITSKTEKEIILLDPLAVKFMSSGFHW PLAAMFESSGQLIAPNDIITVSPR" 6377. .6505 /gene="psaJ" 6377. .6505 /gene="psaJ" /codon_start=1 /product="Photosystem I reaction centre subunit IX" /protein_id="AAC08080.1" /db_xref="GI:1276660" /translation="MNNFTKYLSTAPVIGVLMWTFAGFTEILNRPDPVLYFY" 6660. .7145 /gene="apcd" 6660. .7145 /gene="apcd" /codon_start=1 /product="Allphycocyanin gamma chain" /protein_id="AAC08081.1" /db_xref="GI:1276661" /translation="MSLVQIITINADDELRYPITIGELQIQNVLTGTSNRIRIATIR DKERELIQASQIOPOLHPEYTAGGNAGSKRSCLADYGVYRLITVYIAGPKD SIETGIIIGVRMYSLSGVIITGMDAIQCLAEASILEMGGDADATISYPPYITIRM S" 7473. .8453 /gene="fabH" 7473. .8453 /gene="fabH" /codon_start=1 /product="beta-keetoacyl-acyl carrier protein synthase III" /protein_id="AAC08082.1" /db_xref="GI:1276662" /translation="MGVHLITGSSVVPNSVENQFEDIITSDHWISRTGIRKRL ABSSTLITLAAEANDALSKASINAEIDILLIATSTPDDFGSASQQAIGATSS TAFDITACSGFIALVLTASQFIQAGSVKVLVAGADTWSRMIDMSRSCILFGGA GAVLIGESSINSILGFLCTDGRINSHQLNMSPSQOFGILTVPKGRYDSIRMMGX EYKRAVPROVPYIVKINDVNSITDEVDMPFIHQANTITLAIATRLSIPLSKMTIN LENYGTSPASIPLAIDENIKERKTIQPGVVVLVAGAGLTWGAIVLKQ" complement (8465. .8536) /product="tRNA-Leu" 8735. .8854 /gene="psbX" 8735. .8854 /gene="psbX" /codon_start=1 /product="Photosystem II protein X (4.1 kD)" /protein_id="AAC08083.1" /db_xref="GI:1276663" /translation="MTPSLSSFLNSLLIGAVIWWPITALLFVSGKDRTRNS" 9115. .9981 /gene="accD" 9115. .9981 /gene="accD" /codon_start=1 /product="acetyl-CoA carboxylase carboxytransferase beta subunit" /protein_id="AAC08084.1" /db_xref="GI:1276664" /translation="MSLSWNP.LKKEOSKAHNKSTKQTIIPDGLMYKCCDGLMYKSK VLKRNIAKVPQCSYHFOASSNERIDQLIDQGSWEPMDDVHLISTDPLGFEDQGLYQRL KQATPFTSQDADVQGTGTMLOQKVLGIMDRFMGSGSVGSKTLRLLEKAOER LPAILICASGAGARMOEGMLSQAKVIGSALEMHKKENLILYSVLTSPTTGCVTASFA MGDGLIADPKALIAVAGRVIEGTIKEDLPDNPQSEVLPFHGFLDLIVPTQIRSK LIQIILHITNNN" 10259. .10750 /gene="psbY" 10259. .10750 /gene="psbY"

CDS

10259..10750
/gene="psbV"
/codon_start=1
/product="cytochrome c550 (oxygen-evolving complex component)"
/protein_id="AAC08085.1"
/db_xref="GI:1276665"
/translation="MLKRSWLAFLUGLTVASVSTIYVAIEHDEATRVPLESRT VLTPEQVKRKLNNSCATFNGGCTFSGATGDSLGATFQKNDIEALVVM KDPSTYDGAESIAELHPSIKSAELPPKRNLTDEDLFTIAGHTLLQPKIVSEKMGGX IYV"
gene
10915..11247

Query Match 30.8%; Score 603.6; DB 8; Length 191028;
Best Local Similarity 60.5%; Pred. No. 6,4e-165;
Matches 1029; Conservative 0; Mismatches 669; Indels 4; Gaps 2;

QY 132 GCAACCGGGCTTATTCCTGATGATAGCTGMAACGCCATGGGGTCAACACATTTT 191
DB 118442 GAAACGTGCTGCTTCCTCTTTAGATAGATGTTGTCACAGGATATTCATATAT 118301
QY 192 TGGCTATCCCGCGGGGCAATTTGGCCATCTATGATGAACTGACCGCTTGAACGGC 251
DB 118302 TGGTACCTGGTGGGCTATTTTACCAATTTATGATGAACCTTATGCTTGGAGAACT 118361
QY 252 GGGGGAATTGAGCATTTTGTGCGCCATGAACAAGAGCTTCCCATGCGGAGATG 311
DB 118362 TTCAATTAACAAGATTTCTGTGAACATGAACAAGGCTTCAATGCTGCAATG 118421
QY 312 GTATGCCAGCCCAAGTAAAGTGGAGTTGTTGCTGATCAATCTGACCAAGGCGAC 371
DB 118422 TTATTCAGATCGACAGATGAAGTGTGATGTTTCCCATCTGGGCGTGGTCCAC 118481
QY 372 TAATCTGATGACCGGCAATGCCATGCGCATTTGAGTCCGTCGCTGCTGATTAAC 431
DB 118482 TAAATTAATATCGGCAATGCTAGTACCTGCTGCTGCTGCTGCTGCTGCTGCT 118541
QY 432 TGGAGAGTGGCGGCTGATGATGTTGATGATGCTTCCAGAAATGACATTTTGG 491
DB 118542 TGGACAGTGGAGACCTTTTATGAGAACATGATGATTTCAAGAAATGATTTTCCG 118601
QY 492 CATACCTTACCATGCTTGAAGCACTGCTATGATGATGATGATGATGATGATGATG 551
DB 118602 CATTAATTAATCTTATTAATAATCTATATGATGATGATGATGATGATGATGATG 118661
QY 552 CATTTACTGAGGCTTTCATCTTGTGACACCGGTCGTCGCGGCGCGGTTTGAATGCA 611
DB 118662 AATTTGCTGAGACGTTTATCAATTTGCAAAACATGATGATGATGATGATGATG 118721
QY 612 TATTCCTCAAGATGTGGGCTTGAAGAAATGATGATGATGATGATGATGATGATG 671
DB 118722 TGATCCAAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 118781
QY 672 TAAATCTACCGGCTATGCGCCACGTTAAAGTAAATCCCGCAAAATTAATGCGCAT 731
DB 118782 TAAATCCCTGCTGACAGACCTGATCTAATAATCCCGCAAAATTAATGCTGCTG 118841
QY 732 GCAATTTGTTGAGCAGGCGCAAAATCCCTTGTCTGATGATGATGATGATGATGAT 791
DB 118842 AAAAATATTCACAAATCAAGTACGATTAATTAATGATGATGATGATGATGAT 118901
QY 792 CAATGCCCATGCGCAGGATGACGAATTTGCGAAAGGTTCCAGTTGCGGTAACAACAC 851
DB 118902 TGAATGCTCAAGATTAATTAAGAAATTAATGATGATTAATAATAATCAATCAAC 118961
QY 852 CTTGATGGGAATTTGGGCTTTTGAAGAAACATCCCTTTCGTTGGGATGATGATGAT 911
DB 118962 TCTATATGGCAAGATTTCTTTAATGATGATGATGATGATGATGATGATGATGAT 119021
QY 912 GCATGGCAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
DB 119022 GCATGGGACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119081

QY 972 GGCCCGTTTCACACACCGGCTTATCTGAGCAATTTGCTAGCCGCGCAAGT 1031
DB 119082 AGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119141
QY 1032 AATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
DB 119142 TATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119201
QY 1092 TGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
DB 119202 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119258
QY 1152 CCCACCCATCCCATACACCCAGGCGATGTTAATGATGATGATGATGATGATGAT 1211
DB 119259 TAACTTATTCAGAGCAAAATTAATCTTGGCAAGAAATTAATGATGATGAT 119318
QY 1212 TTACCCCTCCAGTGGCCCATGATGATGATGATGATGATGATGATGATGATGAT 1271
DB 119319 ATACCTCTTTAGTACCAAAATTAATCAATGATGATGATGATGATGATGATGAT 119378
QY 1272 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1331
DB 119379 AACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119438
QY 1332 GCGGCGCATGTTTGAACAAATGACCCCGCGATGATGATGATGATGATGATGAT 1391
DB 119439 GTACGCTCAATTTCTTAAAGTAAATTAATCAATGATGATGATGATGATGATGAT 119498
QY 1392 GATGGCTTTGTTTACCTGCGCATGATGATGATGATGATGATGATGATGATGAT 1450
DB 119499 AATGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119558
QY 1451 CATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1510
DB 119559 AATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119618
QY 1511 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1570
DB 119619 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119678
QY 1571 TCAGTGCACAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1630
DB 119679 ACAGTGCACAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 119738
QY 1631 GCCAATTAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1690
DB 119739 TCTTAATTTTCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGAT 119798
QY 1691 GGAATTTTGGCCCGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1750
DB 119799 ACBAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119858
QY 1751 TGTGCTGCTCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1810
DB 119859 TTGCCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119918
QY 1811 CCAATGCTAGTTTACCGGA 1832
DB 119919 ACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119940

RESULT 12
AB002583/12
LOCUS AB002583 149987 bp DNA circular PLN 07-MAY-2003
DEFINITION Cyanidioschyzon merolae chloroplast DNA, complete genome,
strain:110.
ACCESSION AB002583
VERSION AB002583.1 GI:30409147
KEYWORDS chloroplast Cyanidioschyzon merolae
SOURCE Cyanidioschyzon merolae
ORGANISM Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.
REFERENCE
1

AUTHORS Ohta,N., Matsuzaki,M., Misumi,O., Miyagishima,S., Nozaki,H., Tanaka,K., Shin-I,T., Kohara,Y. and Kuroiwa,T.
TITLE Complete Sequence and Analysis of the Plastid Genome of the Unicellular Red Alga *Cyanidioschyzon merolae*
JOURNAL DNA Res. 10, 67-77 (2003)
REFERENCE 2 (bases 1 to 149987)
AUTHORS Ohta,N., Matsuzaki,M., Tanaka,K., Shin-I,T., Kohara,Y. and Kuroiwa,T.
TITLE Direct Submission
JOURNAL Submitted (02-Apr-1997) Niji Ohta, Saitama University, Department of Molecular Biology and Biochemistry; 255 Shimo-ohkubo, Saitama, Saitama 338-8570, Japan (E-mail:niji@molbiol.saitama-u.ac.jp, Tel:81-48-858-3848, Fax:81-48-858-3384)
FEATURES
source
1. .149987
/organism="Cyanidioschyzon merolae"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/strain="10D"
/db_xref="taxon:45157"
1. .1398
/gene="ycf24"
1. .1398
/gene="ycf24"
/codon_start=1
/transl_table=1
/product="ABC transporter for iron-sulfur cluster formation (sufB)"
/protein_id="BAC76097.1"
/db_xref="GI:30409148"
/translation="MNQQLNLDRLMNLREYVGFSTSPDEYPPFPIGLNEQIVKMSAKCEPOMDIFRLKAVHKKOMPEQMARVARETINFEOLLYAAAPSTEDDYQSLATLGVKSSKVALDPVDSVSTIGTFHQALKDAGVITCSSEALQKPSIVKYYKLYAVYIGDNFSAIANSVPEDEGFCYTPDTCAMELSTFRLOKSGSGQFERTLITAEKASVSYLSGCTAARYDRHQLAAVELIALDLASIKYSTVONWVSQDEGKGKLYNFTKGLCAGEAGKISWTQVETGSAITWKTPSCLANNRADYVSVALTDGYQADTQSKMLH LGANSKR11SKGKICGKSGKNTYRGQVMTSAQKAWNSQCSLLIGKATANTYPIEIANVSARVHEATTTQIGEEQLFYIQQRGINQDAVNWGLGFCQGVFLQPMERA TEVNEELTKLEKNC"
1399. .2139
/gene="ycf16"
1399. .2139
/gene="ycf16"
/codon_start=1
/transl_table=1
/product="ATP-binding subunit of ABC transporter for iron-sulfur cluster formation (sufC)"
/protein_id="BAC76098.1"
/db_xref="GI:30409149"
/translation="MLKIEQLKGIKDRLLIEGDAHEAGSVHAIMGNGSGKSTLAKVIAAPAYKILIEGFWQDESINDKAEARAGLIFLAFQNPPLTIPGVNMDLMLA YKTKKOSIDAISFLEYIMPELVLGADAPFLHRYVNDGSGEKKKEILQALIEPRLALIDEISGLDIALKOIASHMLKTTQNAVVLITHRPALLEYIKPDYVHWQGRRIISGDSASLAKYIERGYDEYKKAAY"
complement (2171. .2275)
/gene="ycf12"
complement (2171. .2275)
/gene="ycf12"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="BAC76099.1"
/db_xref="GI:30409150"
/translation="MNHWVLAQLTVLAFLVAVGPITLIIWLSNKKDALL"
complement (2193. .3065)
/gene="ycf62"
complement (2193. .3065)
/gene="ycf62"
/codon_start=1
/transl_table=1
/product="cell-cycle protein mesJ family"
/protein_id="BAC76100.1"
/db_xref="GI:30409151"
/translation="MYTPMHNKRNRFQLANLTSHSPSHWLAISGGDSLCLKPMDLVHNNHVVVHFDHRSGLAARVYLSRMYHMLAMVFTALPITTEQDARYXANLIDLYOTHSGAICTAHTASDDETYLSDWISLRHPEGIRLGHLSGSLPEPLHLTRAQTHWPTLVHYLPIMSDVSNVOITFKRQHLQHLIPIKQINPFRHTLSARLSARLSPLIVPMHWSLPIWLOQTATWGLTSLQQLIYNKKEILARSSINRFS FCDSSRSYHFNAL"
3086. .3157
/gene="trnK"
3086. .3157
/gene="trnK"
/product="trnK-lys"
/note="codon recognized: AAA"
/anticodon=(pos:3118. .3120,aa:lys)
3183. .3914
/gene="trpA"
3183. .3914
/gene="trpA"
/codon_start=1
/transl_table=1
/product="tryptophan synthase alpha chain"
/protein_id="BAC76101.1"
/db_xref="GI:30409152"
/translation="MLIAYLTAGAPDINTTKEAVMKLAKGADVIEIGVPSDALADG AIIQKASKOALMNGPHLDHMLSEVNEIEVPLVAYNOIMHWGYKWKXIVAH NYKGLIVPDLPIYESKTLQICDRYGINILMISPTNTRQELARACKDWIYISR TGYTGLETSPDKQIPKTLQELKKYTKAPIALGFGISKSEQVRLVSKWADGVITGSAC MQLLEKGYDQSLSEWISVWKSFCP"
3800. .5140
/gene="infB"
3800. .5140
/gene="infB"
/note="might be a partial remnant"
/codon_start=1
/transl_table=1
/product="initiation factor 2"
/protein_id="BAC76102.1"
/db_xref="GI:30409153"
/translation="MGGCGNCKYKCMHNFIRKRSSTIRMDKDEKIIISVGTAVD LAKQKPYTERIRIYAFQCGVIRIINDYDEIMQLQCKGEMQIYVEDTADAEVADARDEDMEIIVASAVEKLEQAPIMSHNHSIAIRKGVYVLOMRKALANLQOITDKDVNRPRAPIIVESRKDNLGIITLLIRKSLKVGDAVIEGNKIRAMQINORPI KEVAGEIVVSGFNQVSLGNTLYSAKKGEKSAQNQIMQKICHPKMLIVASE SESALESGVCFLENNKQLLSAEGSIRKQVLAQTSTSLVAVQTCVMINMNVF PTVKIIKANFEDLLAQLSRYKVRIVVSKALVQVFRQSGIAGCVVIVQGEFKGGE VEIWRNQYQARARLSIETIKHKQAVKAVAPQECGIFKMDTWQVDEIVLYN"
complement (5140. .5212)
/gene="trnT"
complement (5140. .5212)
/gene="trnT"
/product="trnA-Thr"
/note="codon recognized: ACA"
/anticodon=(pos:5177. .5179,aa:thr)
5213. .5785
/gene="trpS4"
5213. .5785
/gene="trpS4"
/codon_start=1
/transl_table=1
/product="30S ribosomal protein S4"
/protein_id="BAC76103.1"
/db_xref="GI:30409154"
/translation="MSRVIQPRVRIIRLGIILPAFTNKSNNRTGVPGEGHAHTRKLS EYAVQLQERKQLQYVYGINNQILARYFRQAKRSRSTGIELKMLERIDHVRAGF APTLPARQDLVNHGVKVNQVITLASFACQVNHIIIEVAKSPSSAQLPPYQVHQNF VKMIQVEKMDIAPRVNELVVEYTRVGA"
complement (5788. .6561)
/gene="ycf80"
complement (5788. .6561)
/gene="ycf80"
/codon_start=1
/transl_table=1
/product="ORF257"
/protein_id="BAC76104.1"

```

/db_xref="GI:30409155"
/translation="MMDYIYERLDEDEKRNOLLQWFVSPVELEHLLSROLOSV
CISEKPSFIQDYQSYSPKNTOLIVQNRASLEACPYYVNPQOEILLSPRWE
LEKAPETLMMWFPHIPIRNPSTSYLAFFPDADAHMYLKSIVNPIERKNSPVEGD
LSEFLKNTKTDIKVILVPIRLOLCKMSKEFGTVPFVQMDHEMLFFSKMDAQOIFKN
QLSKTIYDVISLESVLESDLDQARCTVPLSLNOSL"
gene
complement (6539..6766)
/gene="rp128"
complement (6539..6766)
/gene="rp128"
CDS
complement (6539..6766)
/gene="rp128"
/translation="MSRVCPLTKCNHAEVTVSHKRNHRQFVNLQKKLFINNW
VVLKISTKAIKTLKSKTKSTFNSNALLHL"
gene
complement (6763..7071)
/gene="trxm"
complement (6763..7071)
CDS
complement (6763..7071)
/gene="trxm"
/codon_start=1
/translation="thioredoxin type m"
/protein_id="BAC76106.1"
/db_xref="GI:30409157"
/translation="MLHIDLRENEVLQSEKLVYDFNAPMGPCRMGLIEETAK
ERNLKVQVNTDENPLNLAIFYGIRSLFTLMLFKGGRVDIVIGAVPKSLIHITNKYL
"
gene
7504..8970
/gene="rbcl"
7504..8970
/gene="rbcl"
/codon_start=1
/product="ribulose biphosphate carboxylase large chain"
/protein_id="BAC76107.1"
/db_xref="GI:30409158"
/translation="MAQSVQERTLNKRYESGVIPYAKMYNPNVYVKDSMLALF
RVPQGVDPDIEAABAAGSESTATWVTDLLTACDVRAKAYVDVPAPODYF
VLYADIDLEEGSINLTAITIGNGVFAVKALILEMREIPFSLTKTPOGPAQTVI
VERELKRFEPPLIGCTVYKRLGSGNRYRVYBGLKGLDPLKDEINISPPFRM
RDRLYVBEVNRPAASGEVKSGLYNATVETCYKRAEPAKGVSVIIMDLVIG
YTAIQMAIWARENMIHLHRAGNSTYSQKNHGINFRISIKMMAGVDIHAGTV
Query Match 30.3%; Score 592.8; DB 8; Length 149987;
Best Local Similarity 60.7%; Pred. No. 9,2e-162;
Matches 1030; Conservative 0; Mismatches 652; Indels 16; Gaps 3;

```

```

QY 466 ACCTTACGATGCTTAAAGCACTCCTATGTGTACGTAGTGGCGGCAATATGCTGCATT 555
DB 80051 ACTTGCCATTATGTTAAACATCTTATCTTGTACAGATGCTCGAGATATCTTCGAT 79992
QY 556 GTTACTAGGCTTTCATCTGTTAGACACGGGTGTCGCCGGGCGGTTTATGATGAT 615
DB 79991 ATTGTGAGCCTTACATGCTTAACATGCTGTCTGCTGGCCAGCTTATGATGAT 79932
QY 616 CCGAAGATGTGGCTTAAAGAAATGTAGTACATTCCTCCGACGCCCGGTGACCTTAAT 675
DB 79931 CTTAAGATGTGGGATTAGAAAGAAATGATTAACAGATGCTGTCTCA-----AGTCAAT 79878
QY 676 CTACCGGCTTATTCGCCGCCGCTTAAAGGTATCCCGACAAATTAATGGCGCATTCGA 735
DB 79877 ATCTGTTGTGCGCCCTTATTTTACCACTCTTAACAAATTTAGCAAAATGCTATG 79818
QY 736 TTGTGAGACAGCCAGAAATCCCTTGCTACGTAGGGGAGGGCGATCCGCCAAT 795
DB 79817 ATGTGCAAGAGAGTGTCAACTTTACTTACGTGGGTGAGAGTGCACCAAGC-----T 79764
QY 736 GGCCTACGCGCAGGTGCAAGATTTTGGGAAAGTTCCAGTTGCCGCTAACCAACCCCTG 855
DB 79763 GCTCTCAAGAAATGTATCAATTAAGTACTCAATTTCAAAATCTGTAAACCACTAACCTTA 79704
QY 856 ATGGAAATTTGGGCTTTTGAAGAAACCACCTTCGTGGGTATGTTGGTATGCAT 915
DB 79703 ATGGAAAGATGTTGTTTGAAGAAACGATCTCTTATTTGGGTATGCTTGCATGCAT 79644
QY 916 GGCCACCCCTTATGCCAATCTTCCGTCAGCAAGATGTATTTGATTTGCAAGTGGGGCC 975
DB 79643 GGTACGGCTTATGCCAATTTGCTGTAGCAAGATGTATCTTTAATTTGCTTATGAGTGA 79584
QY 976 GCTTTCAGACGACCGGGTACTGGCAAACTAGACCAATTTGCTTACCGCGCAAGTAT 1035
DB 79583 CGCTTGTATGATCGCGTACCGGTAACTTATGATTAATTTGCTTCTGCAAGATTAAT 79524
QY 1036 CACATTTGATCGACCCGCGAGGTGGGAAAAACAGGGCTCCGATGTCCTCATTTGTG 1095
DB 79523 CATATGATATTTGATGCTGCTGAATTCGAAAAAATGGAATTCCTCATTAAGCATCTGT 79464
QY 1096 GGGATGTAGACCATGTTTGAACAGCTTTGACAGGGGGCGGGAATGATTAACCC 1155
DB 79463 TCTGATGTCAAACTTGTTTAAACAATTTCTTACCTGCACTTAATTTGCTCTTAATTT 79404
QY 1156 ACCATCCCATACCAACCCAGCATGTTAATGCAATTCATTTGGCGGACCGATTAAC 1215
DB 79403 GTTGCTATGATTAAGAACTATGCTTGGAGAAAGATATGACGTTGAAACAGATTAAT 79344
QY 1216 CCCCCTCAGTGGCCCACTATGAGATATTAATGCCCCCAGAGATGATTAAGAAAT 1275
DB 79343 CCTTACTAATTCCTGCTCTCTCTTATCTATCCGACAAAGAAATTTAAATGCACTT 79284
QY 1276 GGTGCGCAGGCCCCGATGCTTACTACACACCGATGTGGGACAAACCAATGTGGGCG 1335
DB 79283 GGCCTAATTAAGCTCTCATGCTATGATTAAGTACGATGAGACAAACATTAATGTGGCA 79224
QY 1336 GCCCATGTTTGAACATGATGCCCCCGCGATGATTTTCAGTGTGCTTGGGTGATGATG 1395
DB 79223 GCACAGTTATTAAGGTTCACACCCGTCMAAGGCTTTCACGCTTGCTTAAGTACATG 79164
QY 1396 GACTTGTGTTTAACTGCGCCGATGAGGACCAAGTGGAGATGAGGAGAGAGCGGTG---- 1451
DB 79163 GGCCTATGCTTACCTGACGCTATAGTGTGCTGCTGCAATCAAGCAATGATTCAT 79104
QY 1452 ATTTCATCAATGATGATGCAAGCTTCAATTAATCTTCAAGAACTGGGAAACCTTAGCC 1511
DB 79103 ATTTCATTTCTGATGATGCAAGTTTCAATTAATTAATTAAGAACTGACCAATTTCA 79044
QY 1512 CAGTACGACATCCAGTTAAAGTAACTATTTCTCAATAAGTGTGGCAGGGGATGAGTGT 1571
DB 79043 CAAATTAAGCTCAAAATTAATTAATTTCTTTTCAAGATGTGGCAAGATATGTAGCT 78984
QY 1572 CAGTGGCAACAACTTTCTACAGAAAGCTTATTTCTGCTTAAACATGTCCAGAGGATG 1631

```



```
gene      /anticodon=(pos:50..52,aa:Tyr)
           /complement(181..573)
           /gene="rp119"
           /complement(181..573)
CDS        /gene="rp119"
           /codon_start=1
           /product="ribosomal protein L19"
           /protein_id="AAC35594.1"
           /db_xref="GI:3602933"
gene      /translation="MEIILHLYEFVNMKSNKLIQLVESHNSKTLFQLQVGDVRLG
           VSITGEKRIQATEGVITKSNCGINLTITRRVLQIGIERVYLHSPKLSIEIL
           RKSKIRAKLYIARTSGKATRLQGRFN"
           729..3188
           /gene="clpC"
           /gene="clpC"
           729..3188
           /gene="clpC"
           /codon_start=1
           /product="Clp protease ATP binding subunit"
           /protein_id="AAC35595.1"
           /db_xref="GI:3602934"
           /translation="MPEPTEKAIKYIMLAQSEARLGHNPVGTBOIILIGEGTGI
           AARVLSGMVNLKDAVEVKILIGSGFVAEIPTPPAKRVLELSBEARQLGHNH
           IGTEHLLGIRGEVAAVLENLADLTAVTQVIRLLGDAEVSATNGTKKTP
           TLEPFGSNLTQKAEGKLPVIGROKEIERVIOILGRKKNPILIGEGVAKTIAE
           GLAQRINRDPDILEDKRVTLDIGLVAGTKYGEPEELKIIDERVANNILV
           IDEVHTLIGAAEGAIDANILKPLALGEMOCIGATLLEERKIIKDSALERRFO
           PVYGEPSVRETIELIYGLADRYEKHKVILSDPAISAAPADQYIADRPLDKAID
           LIDEGSRVRLNMSQLPAPARELDKELIKOKDAVNSQDPETAGQLRDREMIKA
           QIAIAHSKGDENITEVSATVEEDIAQVAATGI PVNKMTRSEKLDLQMEELH
           GRIGDEAAVAVSAKIRARVGLKNPNRIASFISGTVGKTELTAALSPBEIGH
           BEAMVRLDMEYMERHVSRLIGSPGYVNGGQLTSVRPRPYTVLPBIEKHG
           PDVENLLLOILEGRLDSSKGTVPDNTLLITSVNGSKVIEKGSGGAGPSEIQT
           ESQYGRKALVNEIKOYPRPELNRLDEITVROLKDEVGELAIMKKEVTRISE
           KGIOLETTARFKHLINEGNPIYGARPLRARVRLLEDTLSEPLAEIKIKGDTAVV
           DVDDGKVKVLFSEKLELVAN"
           3316..3388
           /gene="trnF(gaa)"
           3316..3388
           /gene="trnF(gaa)"
           /product="trnF-Phe"
           /note="codon recognized: UUC"
           /anticodon=(pos:3349..3351,aa:Phe)
           /complement(3402..5201)
           /gene="dnaB"
           /complement(3402..5201)
CDS        /gene="dnaB"
           /complement(3402..5201)
           /codon_start=1
           /translation="MKVTKEFYNLKAEMHLLIQLIVKEDFLQITRKIDTQVYFQSHQ
           ILVRSNIHLTKKVLINLPMILSHSELVSEIVAVLITLPMRDNPRILINDSLK
           VLIDNFIRELQNSCAKIVSLTFNENISVETLLQKSNLISINTKKVSLSOL
           LLLETILEIDDKTNRSHVLTGFEDLIHVLGLOKSLIITAGPSMGKATAPMISIVRN
           VADIQSPPIVIFLSMSSKQIIRLISRENIATSRLENISISEMELINRAMTILS
           NLNLTVIDDENLVDLQSKLASLQOYIGDILVIDYQLDQYKESQSRITELSYI
           TRYKLTAKDFNLPLVVLSQLSRVFRILKRPILSDKESGCLSYATQOYFLKSDN
           VNFEKLSLKSNHYISATIELIPQVARIYPIVSLIRIEQTKYKVVLEDPYQ
           VIVSTEYVYLLNGMKRIELFTVDLVALDIOYLIYNNTEVDLFSNVIFESVNL
           ICMNRIVVDPMI PKNTPFVNALIVNSI EODADVLFYRDSYVGHMLKQNMIDMC
           EYIYAKRHGHTIGMNLINPNTVTSFNNLIKES"
           3594..4073
           /note="signal element intein; encodes putative intein
           spliced out of the replication helicase subunit"
misc_feature
           5361..5433
           /gene="trnM(cca)"
           5361..5433
           /gene="trnM(cca)"
           /product="trnM-Tip"
           /note="codon recognized: UCG"
           /anticodon=(pos:5394..5396,aa:Tyr)
           5588..6013
gene
```

```
CDS      /gene="rp111"
          5588..6013
          /gene="rp111"

Query Match      30.3%; Score 592.6; DB 8; Length 121524;
Best Local Similarity 60.2%; Pred. No. 1e-161;
Matches 1018; Conservative 0; Mismatches 669; Indels 4; Gaps 2;

QY      136 ACCGGGGCTTATATCCGATGATGATAGCTGAAAGCCATGGGGTCAAAACATTTTGGC 195
DB      119823 ACAGGTGATTCGCAATTAATGATAGTATAGTTAGTTGCGCATGATGTAAACATATTTTGGT 119882
QY      196 TATCCCGCGGGCAATTTTGGCCATCTATGATGAATCTATGACGCTTTGAAGCGCGGG 255
DB      119883 TATCCAGGTGTCATTTTATACCTATATATGATGAATCTTATGTTGGGAAAAAGAAAGA 119942
QY      256 GAAATTAAGCATTTTGGTGGCCATGAACAGAGACCTTCCATGGCGCGGATGGGTAT 315
DB      119943 TTATATGAGCATATTTAGTTGCGCATGAACAGAGACCCCTCATGTGATGACGATAT 120002
QY      316 GCCAGAGCCACAGGTAAGTGGAGTTGTTGGTACATCTGACACCGGGCGACTAAC 375
DB      120003 GCAAGATTAACAGGTAAGCTAGTGTGTGTGTTTGGCTACTTCAGGACCCGGTGTCAAAAT 120062
QY      376 TTGCTGACCGGCATTCGCAATGCCATTTGGACTCGGTGCCCATGTGTGTGATTA 435
DB      120063 TTAGTAACAGGATATGCTACTGCTCATATGGAATCAATGACGCTCATTAATTAACGGGT 120122
QY      436 GAGTGGGCGCTGCATGATTTGTAAGGATGCTTCCAGAAATTAATGATTTTGGCATC 495
DB      120123 CAAGTAGGACGCTCTTTATAGGTAAGTATGATTCGCAAGAGTGCATATTTTGGTAT 120182
QY      496 ACCTTACGATGCTTTAAGCACTCCATGATGATGATGATGATGATGATGATGATGAT 555
DB      120183 ACTTACTGATTTGTTAAACATCATATGTTGTCGTGAAACAAAGATGGGTAAATA 120242
QY      556 GTTACTGAGCTTTCATCTTGTCTAGACACCGGTGTCCCGGCGGTTTGTATGATAT 615
DB      120243 GTAGCAGATCATTTTATATAGCAAAATATAGTGCACAGGCTCTGACTAATATGATATA 120302
QY      616 CCCAAGATGAGGCTTGAAGATATGATGATGATGATGATGATGATGATGATGATGAT 675
DB      120303 CCMAAGATGTTGGGCTGAGAAATTTGATTTATCAATTTGTAATCCAAATATATATAAC 120362
QY      676 CTACCGGCTTATCGCCCCACGCTTAAAGTATATCCCGCAAAATTAATCGCGCAT 735
DB      120363 TTACGAGATGCTCTGCTTAAAAAATATGATCAGATATGATATGATCAAGAGCTAAT 120422
QY      736 TTGTTGAGCAGCGCCAGAAATCCCTGTCTACGAGGAGGAGGCGCATGCGCCCAAT 795
DB      120423 TTAATTAAGCAATCCAGCAACCTTTATTAATATATGATGATGATGATGATGATGAT 120482
QY      796 GCCCATGCCAGGTGCAGAAATTTGCGGAAAGTTCCAGTTGCGGTTAACCAACCACTCG 855
DB      120483 TCACACACGAATTAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 120542
QY      856 ATGGAATATGGGCTTTGAAGAAACCATGCCCTTCCGGGTATGTTGGGTATGATCAT 915
DB      120543 ATGGGTAAAGAAATTAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 120602
QY      916 GGCCACCCCTATGCCCACTTTGCCGTGACGCAATGATTTGTTGATTTGCAAGTGGGCGC 975
DB      120603 GGTACTGTTTATGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 120662
QY      976 CGTTTGAAGACCGGGTATACGCAACTAAGCAATTTGTTAGCCGCGCAAGTATTT 1035
DB      120663 CGCTTTGACGACCGGTGTAACAGTAAAGTAAAGCAATTTGCTGTCATGCAACAAAGTATTT 120722
QY      1036 CACATTGACATGACCCGCGGAGGTGGGAAAAAACAAGGGTCTCCGATGTGCCATTTGTG 1095
DB      120723 CATGTAGATATGATCCAGCTGAATTTGGAGAAATGCAACCCCAATGCGTATTTGTA 120782
QY      1096 GGGAGTGTACGCGCATGTTTGAACACGCTTTTGAAGCGGGCGCGGAAATGGAATTACCC 1155
```

Db 120783 GG---TGAATCAAAAGACTTTTCCGTGATTTAATGTAATGTTAAAAAATACATTAT 120839
 QY 1156 ACCCATCCCATACACACCCAGGAGCATGTTAAATCCCATGATCATTTGGCGGACGATTAC 1215
 Db 120840 TTGAGACGAGACAGAGCCCACTTGGAGTCAACCAATATAGTGGGAAAAAGAAAT 120899
 QY 1216 CCCCCTCCAGTGGCCCACTATAGATTAATTTGCCCCCCCAGAGAGTATGACACGAAAT 1275
 Db 120900 CCACCTCTTGTACAAAATAATTAATTAATTAATCAACAAGAAAGTTATTCACGAAAT 120959
 QY 1276 GGTCCGACAGCCCGGATGCTACTACACACCGATGAGGAGCAACACCAATATGAGGCG 1335
 Db 120960 AGTACAGAGCTACAAATGCTTATTTACACAGATGTTAGCCACATCAATATGAGGCG 121019
 QY 1336 GCCCAGTTTGAACAAATGAGCCCGCCGAGATTTCCAGTGTGGCTTGGATACGATG 1395
 Db 121020 GCTCAATTTATCAAAAAGTCAAAAAGCATGATTAACAAGTGTGGTTAGAACATG 121079
 QY 1396 GGCTTGGTTACCTGCGCGCATGGAGACCAAGTGGAGAGTGGGGAGACGAGC-GGTCAAT 1454
 Db 121080 GGAATCGGTTACCCGACGCTATAGGTGACAAATAGCACATCCAAATGAACAAGTTAT 121139
 QY 1455 TGCAATCAGTGAAGATGCCAGCTTCCAAATGATCTTCAGAACTGGAAACCTAGCCGAG 1514
 Db 121140 TGTATCTGTGTGATGCAAGTTCCAAATGAATATCCAAAGACTAGTACAGTATACAA 121199
 QY 1515 TACGACATCCAGGTTAAATCTATTTCTCAATTAAGGTTGGCAGGGAGTGGTGGTCAG 1574
 Db 121200 TATGATTAACCAATTAATAATTTATTTATTAATAATTAATGCAAGGATGTAAAGCAA 121259
 QY 1575 TGGCAACAACCTTTCTACGAAAGCGTTATCTGCTTCAATGATCCAGGGCATGCA 1634
 Db 121260 TGGCAGCAAGCGTTTACGCTGAAAGATTTACATTTAATATGAAAAAGTGCTCTCT 121319
 QY 1635 GACATTAATCTCCCTCTGTGAAGCCTATAGGCATCAAGGGTATTAATCTGTCGCAAGCGGGAA 1694
 Db 121320 AATTTTACGAAAGTGGCGAAGCATTTGGCTTACGCTCACTGAAATTAATCCCGTAC 121379
 QY 1695 GATTGGCGCCGCGGATGGCGAAATGCTAGCCCATGATGCTGTGTGTATGATGATG 1754
 Db 121380 GATTCCAACTGAGATTAAGAGGCTTTAGATTATGATGACCTATACCTAGTGAATA 121439
 QY 1755 GTGTCAAAAAAGATGAANAACGTTACCCCTATGATTTGCCCGGATGATTAATGCCAA 1814
 Db 121440 CAAGTTATCGCAGATGAANAAGCTTATCCATGTTGCTCCAGGAAAGATTAATGCACA 121499
 QY 1815 ATGCTAGGTTT 1825
 Db 121500 ATGATGGTAT 121510

RESULT 14
 AF022186 164921 bp DNA circular PLN 14-DEC-2000
 LOCUS Cyanidium caldarium strain Rkl chloroplast, complete genome.
 DEFINITION AF022186 236235 270297
 ACCESSION AF022186.2 GI:6466296
 VERSION
 KEYWORDS
 SOURCE chloroplast Cyanidium caldarium
 ORGANISM Cyanidium caldarium
 Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales;
 Porphyridiaceae; Cyanidium.
 REFERENCE 1 (bases 130696 to 132364)
 Vogell,H., Fischer,S. and Valentin,K.
 TITLE A model for the evolution of the plastid sec apparatus inferred
 from secY gene phylogeny
 JOURNAL Plant Mol. Biol. 32 (4), 685-692 (1996)
 MEDLINE 97134960
 PUBMED 8980520
 REFERENCE 2 (bases 1 to 164921)
 Gloeckner,G., Rosenthal,A. and Valentin,K.
 TITLE The structure and gene repertoire of an ancient red algal plastid

genome
 JOURNAL J. Mol. Evol. 51 (4), 382-390 (2000)
 MEDLINE 20496959
 PUBMED 11040290
 REFERENCE 3 (bases 46857 to 47851)
 AUTHORS Valentin,K.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-1996) Institute for Plant Physiology, Justus
 Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany
 4 (bases 28701 to 75580)
 REFERENCE Gloeckner,G., Rosenthal,A. and Valentin,K.
 AUTHORS Direct Submission
 JOURNAL Submitted (02-SEP-1997) Department of Genome Analysis, IMB Jena,
 Beutenbergstr.11, Jena 07745, Germany
 5 (bases 1 to 164921)
 REFERENCE Gloeckner,G., Rosenthal,A. and Valentin,K.
 AUTHORS Direct Submission
 JOURNAL Submitted (18-NOV-1999) Genome Analysis, Institute for Molecular
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
 6 (bases 130696 to 132364)
 REFERENCE Vogell,H., Fischer,S. and Valentin,K.
 AUTHORS Submitted (18-NOV-1999) Institute for Plant Physiology, Justus
 Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany
 On or before Nov '93, 1993 this sequence version replaced gi:529651,
 gi:1240002, gi:2465730.
 FEATURES
 Location/Qualifiers
 1..164921
 /organism="Cyanidium caldarium"
 /organelle="plastid:chloroplast"
 /mol_type="genomic DNA"
 /strain="Rkl"
 /DB_xref="taxon:2771"
 /join(164808..164921,1..78,462..644)
 /note="direct repeat"
 /repeat_unit separated by stem loop"
 339..477
 /note="stem-loop separates direct repeat unit"
 580..780
 /note="similar to yeast mitochondrial origin of
 replication region in Genbank Accession Number J26902"
 986..2041
 /gene="psbD"
 986..2041
 /gene="psbD"
 /product="unknown"
 986..2041
 /gene="psbD"
 /note="Photosystem II D2 protein"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAF13026.1"
 /db_xref="GI:6466445"
 /translation="MTIAGREQERGWPDLLDMDIKRDFEFGSGILLPPCAVIAL
 GAWFTGTTVSSWYTHGLASYLEGNGEPLTAIVASBPANMGSLFLFMPEAQCFTR
 WCQIGLWFTALHSGPGLGICLROFETARLVGAPRYAIAFSGIVAFVFLPY
 LQASWEPAPSPFVAIAIPRLLFLQGFHWNTLPPHMGVAGILLGALLCAHGAFTV
 NTLFENGASDTPRAFTPROSEETYSMTANRPMQIIGVAFANKRMWLFLLFPVPT
 GLWVSGTIGVALNIRAYDFVQSIRADEPFEFTYKILLNKGIRAWAADQDP
 HENFVFEEDLPRGNAL"
 2031..3413
 /gene="psbC"
 2031..3413
 /gene="psbC"
 /product="unknown"
 2031..3413
 /gene="psbC"
 /note="Photosystem II 44 Kd apoprotein"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAF13025.1"

```

/db_xref="GI:6466444"
/translation="MLFNKDKNNGISRESTGFAMWSGNARLINLSGKLLGAHVAH
GLIVFTGAMTLEPSTHPIREKPLYESQMILPHLATLGWVAGGEIVNTYPTFG
VITHVSAUVAFGGIVHSIVGPDLTDFSPFPCVDMRDKKMTTIGIHLILIGAF
LIVKALFTGGIYDITAPGGADIRFTINPLNPAIIIPSYLXSPFGSEEMIVGNME
DVIIGHIWTGIVTIGIWHILTRPESWARRAVSGEATLSTLGALMLGQTALEY
AWNNIVYSEFPGPTAAEASQAQAPFLVDRQLGANLSTGPGTGLKYLMSRSTG
EVLIGETRMFMDLRAPMLEPLRSSNGLDNKIKNDIQPWERRAAEVMTHTPLGSLN
SVGVATEINSVNVSPRSWLTSTSHFPGFIFIGHLMHGRARAAMAAAFKGINREN
EPTLSMRPLD"
3546..3821
/gene="ycf19"
/note="similar to ycf19 in Porphyra purpurea"
3546..3821
/gene="ycf19"
/product="unknown"
3546..3821
/gene="ycf19"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3024.1"
/db_xref="GI:6466443"
/translation="MKDQDLETTTAPQIYIYIILMSLGMFPENINWYSQFYS
LSQSDPYNLPHGVPPSPFGIDSPPIGTLIDFIIEILSRQLKPL"
3873..4133
/gene="ecrf9"
3873..4133
/gene="ecrf9"
/product="unknown"
3873..4133
/gene="ecrf9"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3023.1"
/db_xref="GI:6466442"
/translation="MNNCLPKNKVFXTLNEFKCIDPWCNTDLFKFLYKLNKKKQKT
NKFKYKNTKKKVIDIMLIFYIYHIVIRIKNQTKDQITKC"
4189..4272
/gene="ecrf10"
4189..4272
/gene="ecrf10"
/product="unknown"
4189..4272
/gene="ecrf10"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3022.1"
/db_xref="GI:6466441"
/translation="WTKLKSLLNDKULTLSLUNKTIYDIN"
4334..4406
/product="rRNA-Thr"
4450..4971
/gene="ycf3"
/note="similar to ycf3 in Porphyra purpurea"
4450..4971
/gene="ycf3"
/product="unknown"
4450..4971
/gene="ycf3"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3021.1"
/db_xref="GI:6466440"
/translation="WPRSQRDNPIIDTFTSIADLIYKILPTKESKEAEVYVKDKA
AOSGEVBAALACYQAKTKEDPMDSFTIYNTGLQASNGARLERYHESLKEN
ENLVQALNNIIVYHYGNKLFESKIQEALMPDKASNYWRKAIKLAPNYIEADNW
LKITGRITEDIIML"
5004..6476
/gene="atpb"

```

```

mRNA
5004..6476
/gene="atpb"
/product="unknown"
CDS
5004..6476
/gene="atpb"
/note="ATP synthase CF1 beta chain"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3020.1"
/db_xref="GI:6466439"
/translation="MKPQOLKQKQTLAKVNNKGVILQIGPLDIVEPDGQLPKV
FNALKINNSNNMTTCCEVOILGDNKVRVAMSTTEGLKGGASALDGEPSIIVGKE
TLGRIFNVLGEPIDEKPVINDLPIHRRAPKTCQLEPDSIFENGITKVIDLAPVR
RGKIGLPGAGVAKTVLIMELINNVAHAGVSFVGVEERTREGDVIQEKESGV
INEKDNLNLSYALCYQMNEPFGARMEVGLALTMAEYFRDVANKVNLFTDNTFRV
QAGSEVSLALGRMPSAVGYQPTLIGTEGALQERITSLDQISITSIQAVVPADLTD
APATFPAHLDATVLSRALAKGIYPAVDPDSTIMLOGIYSDEHYTTARKKETL
QRYKELDIIIALIGLDESEDRILISRRARIEKFLQPFVFAVEVFGISGKYVSLSD
SIKGFMTLSGEVNIIPQAFYLVGRIBEAIDKAKQVEKS"
6486..6887
/gene="atpb"
6486..6887
/gene="atpb"
/product="unknown"
6486..6887
/gene="atpb"
/note="ATP synthase CF1 epsilon chain"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAFL3019.1"
/db_xref="GI:6466438"
/translation="MALTVKVTDPDVRVWKKTVVEIILPSITGOLIMNHAPLTL
DIGVRARMNTWVPLVILGFAFDIDNNLVITIIYSDBEKVAIDEBEANKLLASLNL
Query Match 29.8%; Score 584; DB 8; Length 164921;
Best Local Similarity 59.7%; Pred. No. 3,6e-159;
Matches 1019; Conservative 0; Mismatches 680; Indels 7; Gaps 2;
134 AAACCGGGGCTTATTCCTGATGATGATACCGCGAAGCGCATGGGGTCAACACATTTTGG 193
73843 AACCTGGGGGCTTCTGCTTAAATAGATGCTGCGTAAACAAAGTAAATATTTTGG 73902
194 GCTATCCGGCGGGCGCAATTTTGGCCCATCTATGATGAACCTGACGCTTTGAAGCGCGG 253
73903 GATATCCAGAGGAGCGATTTCTCCCATCTATGATGAACCTATCACTGGGAAAAAAGA 73962
254 GGGAAATTGAGCATTTTGGTGGCGCCATGAACAGAGCTTCCCATGGCGGATGGGT 313
73963 AACTAATTAAACATATTAGTTCGTACGAACAAAGTGCAGCTCATGACAGCATGCTT 74022
314 ATGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGTACATCTGACACAGGCGGAGCTA 373
74023 ATGCCAGAGCACTAATGAAGTAGGGGTGTGTTAGTCAKATCTGGGCTGGCGAACA 74082
374 ACTTGGTAGCGGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCA 433
74083 ATTTAGTAACAGAAATAGCTACTGCTCAAGTGAAGTCACTGTTCCATATTTCTATTAG 74142
434 GAGAGTGGGCGCGGCCAATGATGTTGATGAGATGCTTCCAGAAATGACATTTTGGCA 493
74143 GACAAAGTAGCAGAGCATTTATAGTACAGATGCTTTCAAGAAAGTGAATTTTGGAA 74202
494 TCACCTTACCATCGTTAAGCACTGCTATGTGGTACGTAAGTGGCGGATATGCTCGCA 553
74203 TTAACCTACCGATTTAAACATTCATTCGATGATGAAGAGACCCAGAAATATATCACA 74262
554 TTGTTACTAGAGCTTTCATCTGTGTACACCGGCGTCCGCGCGGTTTGGATGATA 613
74263 TCGTTTCAAGACATTCATATTTAGTAAACATGATGAACGAGTGTCTGTTTAAATGATG 74322
614 TTCCAAGGATGTGGGCTTAGAAGATGTGATACATTCCTCCGACACCC-----GGTGG 667

```

Db 74323 TACCTTAAGATGTTGTAAGTAAGTAATTAATCATATGATGATGCTCAATCCGTGATC 74382
Qy 668 ACCTTAATACCGGGTATACGCCACCGTTAAAGTAATCCCGACAAATTAATGCGG 727
Db 74383 ACAACCTTAATTAATACAGGCCCAATATAGCGGCTTGACATTAAGACAAATTTGAAAAAT 74442
Qy 728 CATGCAATTTGTGAGCAGCGCAAAATCCCTGCTTAACGTAGGCGGAGCGCGATCG 787
Db 74443 TCAAAAAATGCTTTTAGAGTCCAAACAGCTTAATGATGTTGAGCGCGGCGCTGTA 74502
Qy 788 CCGCCATGCGCCATGCGCCAGTGCAGAAATTTGCGGAAAGTTCAGATTGCCGATGACAA 847
Db 74503 TGTCTAGAGCTACGACGAAATGAGAAATAGCAAGCTTTATAAAAATTCGGGTACCA 74562
Qy 848 CCAACCTGATGGGAATTTGCGGCTTTTACGAAAAACCATCCCTTCGGTGGATGTTGG 907
Db 74563 CAACCTTAATGGGGAAGGATCTTTTAATGAATACATCTCTTGTATCTTGGAAATGCTTG 74622
Qy 908 GATGATGAGCCACCGCTATGCGCACTTGGCGTACGGAATGTGATTTGTGATGACG 967
Db 74623 GATGATGAGAACTGCTTATAGGAATCTGCTGTAGCGAGTGCATTTGCTCATTTGCT 74682
Qy 968 TGGGGGCGCGTTTCAGACGCGGTAATGCGAACTAGACGAATTTGCTAGCCGCGCA 1027
Db 74683 TAGGGGCAAGATTTGACGACAGAGTGACTGAAAACTAGATTAATTTGATGCAATGCTC 74742
Qy 1028 AAGTAATTTACATGATGATGACATGACCGCGGAGAGTGGAAAAACAGGCGCTCCGATGTC 1087
Db 74743 AAGTAATTTACGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 74802
Qy 1088 CCATTGTGGGGATGATGACCGCATGTTTGAACAGCTTTTTCAGCGGGCGCGGAATGG 1147
Db 74803 CATTAATTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 74862
Qy 1148 ATTATCCCAACCATTCCTCCATACCGAGGATGTTAATGCAATTTGATGATTTGGCGGA 1207
Db 74863 CAAACATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 74922
Qy 1208 CCGATTAACCCCTCCAGGTGCGCCCATAGAGATTAATTTCCCGCCAGAGATGATGAC 1267
Db 74923 AAGATATTCATTAAGATCTCTCTGATGATTAATCTTATATCCAGAGATTAATTA 74982
Qy 1268 ACGAAATTTGTGCGCAGCGCCCGCATGCTTACACGACCGATGCGGACCAACCAAA 1327
Db 74983 ATGAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75042
Qy 1328 TGTGGGCGGCGCGATTTTGAACATGCGCCCGCGATGATTTCCAGTGTGCTGGTGG 1387
Db 75043 TGTGGGCTGCTCAATTTTAAAGTTGAACAGGAAATGCTATCAAGCTCTGCGCTTG 75102
Qy 1388 GTACGATGGGCTTTGTTTACCTGCGCCATGCGGACCAAAAGTGGGAGTGGGCGGACG 1447
Db 75103 GTACGATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 75162
Qy 1448 -AGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1506
Db 75163 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75222
Qy 1507 TAGCCAGATGACATGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1566
Db 75223 TAGCACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 75282
Qy 1567 TGGCGATGTCGACAAATAATTTTCAAGAAAGCTTTATTTGCTTTCAACATGTCGACG 1626
Db 75283 TCAGGCAATGCGCAAGCTTTTTCAGACCGAGGATGCGCACCTTAACATGCGCAAG 75342
Qy 1627 GCATGCGCAGACATTAATCTCTGCTGAGGCTATGATGATGATGATGATGATGATGATG 1686
Db 75343 GACAACTGATTTTGTTCATTAATTAAGAACTCTTACGGGATTAAGAGATTAAGATCA 75402
Qy 1687 AGCGGAAAGATTTGGCGCGGATGCGCGCAATGCTTACCCACATGCTCTGTGTGA 1746

Db 75403 CAAGCAAGATTTAAATCAAAAATTAAGAAATTAATATGACACACTGCGCCCTTTGTTGA 75462
Qy 1747 TGGATGTTGGTGTCAAAAAGATGAAAAAGCTTTACCTATGATGATGATGATGATGATGATG 1806
Db 75463 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75522
Qy 1807 ATGCCCAATGCTAGTTTACCGGAA 1832
Db 75523 ATTCTCAATGTTAGACTCAAAA 75548
RESULT 15
LOCUS AE017162
DEFINITION Prochlorococcus marinus subsp. marinus str. CCMP1375 section 2 of 6 of the complete genome.
ACCESSION AE017162
VERSION AE017162.1 GI:33237292
KEYWORDS
SOURCE Prochlorococcus marinus subsp. marinus str. CCMP1375
ORGANISM Prochlorococcus marinus subsp. marinus str. CCMP1375
REFERENCE 1 (bases 1 to 300355)
AUTHORS Dufresne, A., Salanoubat, M., Paterinsky, F., Artiguenave, F., Kham, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, F., Makarova, K. S., Ostrowski, M., Oztas, S., Robert, C., Rogozin, I. B., Scanlan, D. J., Tandeau de Marsac, N., Weisenbach, J., Winkler, P., Wolf, Y. I. and Hess, W. R.
TITLE Genome sequence of the cyanobacterium Prochlorococcus marinus SS120, a near minimal oxyphototrophic genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (17), 10020-10025 (2003)
REFERENCE 2 (bases 1 to 300355)
AUTHORS Dufresne, A., Salanoubat, M., Paterinsky, F., Artiguenave, F., Akmann, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, F., Makarova, K. S., Ostrowski, M., Oztas, S., Robert, C., Rogozin, I. B., Scanlan, D. J., Tandeau de Marsac, N., Weisenbach, J., Winkler, P., Wolf, Y. I. and Hess, W. R.
TITLE Direct Submission
JOURNAL Submitted (28-May-2003) UMR 7127 CNRS et Universite Pierre et Marie Curie, Station Biologique de Roscoff, Place Georges Teissier BP74, Roscoff 29682, France
FEATURES
source location/Qualifiers
1..300355
/organism="Prochlorococcus marinus subsp. marinus str. CCMP1375"
/mol_type="genomic DNA"
/strain="CCMP1375; SS120"
/sub_species="marinus"
/db_xref="taxon:167539"
86..706
/gene="trxA"
/locus_tag="Pro0315"
86..706
/gene="trxA"
/locus_tag="Pro0315"
/locus_tag="Pro0315"
/codon_start=1
/transl_table=11
/product="thioredoxin family protein"
/protein_id="AA99361.1"
/db_xref="GI:33237293"
/translation="NWKTSITMLPLGYAMPSPDLPLQINDTLFSSSNTAARFTSNM LANKPVLIMICAHCPFNVISLSESLKLDQDGDIDQFLAVASNSVSHPEDEGPEFLV QOIIRKNGAPYLLDLEOTPAKSLKACTPDLFSLPYPKSGSHLLYRQQLDSRRCN NIDPNEGFDLRRLNVLNVCSEVPIDQKPSIGCNIKRSEFSESWIG"
769..1956
/gene="wecE"
/locus_tag="Pro0316"
769..1956
/gene="wecE"
/locus_tag="Pro0316"
/locus_tag="Pro0316"
/codon_start=1
/transl_table=11

/product="DegT/DnrJ/EryC1/Stx5 antibiotic transferase family
enzyme"
/protein_id="AAP9362.1"
/db_xref="GI:33237294"
/translation="MQVPPPTLEKQSLSTIGDDEAVSRVYKSGIYIGSEVLTKEFOS
FADVAVGSPALIGNSGTDAVLALRALNIGPMDEVTTPSPATKAEIASNGATVPS
VADVPYLYLDIELIEAIIPTKAILPVLFGCPVNMTKIVSLAKNGIKYIEDAQ
AVGTQMDGKRVGSGIDIGCFSPFTKMLGAAGAGATTNDPELARIIEHLAIGHPC
RYFTEIGYNSRLDAVOAILNVALPILDKWITQKREIARYINLKGISDLAPFC
EKNDHISMNQFVIRKDCSEKVKGIKIKENIKDLPESDARDLTKRLLHDGVNT
IYYPIPIHQPAKKGKVKYNNLPNTSICQVLSIPPELTLEGEVYNSIKKS
LKD"
/complement(1994..2554)
/locus_tag="Pro0317"
/complement(1994..2554)
/locus_tag="Pro0317"
/codon_start=1
/transl_table=1
/product="NMDX/MutT family pyrophosphatase"
/protein_id="AAP9363.1"
/db_xref="GI:33237295"
/translation="MSPLSQQSPKILNIOALDSKKIREINRFEPLIGVECTFGVI
KHPGALAVPIKENGAVIYROYFAISRIILFEPAGTLEPNEAPLKTIORELOESG
YKASKMDSLVGVMPCPGVSDIILHYLARETEIPILKQSGDEDEDIEVLTMTSELD
CIASGNALDSKSVTAMHRACOVLK"
/complement(2626..3207)
/gene="FolK"
/locus_tag="Pro0318"
/complement(2626..3207)
/locus_tag="FolK"
/codon_start=1
/transl_table=1
/product="7, 8-dihydro-6-hydroxymethylpterin-pyrophosphokinase"
/protein_id="AAP9364.1"
/db_xref="GI:33237296"
/translation="MSSINLYISKTLAIGGANIESPAGSPAKTLITAPRLNITTS
WIKSKSNASNSTEDSTIDPSMLARTALIGGPKTOPFINALVIRGGINOLIP
SVKARDLNLLILKFRGRDRISIKIHNGPRSLDILAMGDLQINBSDLILPHPY
LIERDPVILPLAQLVKSSSPIQIPQKQKPE"
/complement(3255..5378)
/gene="chld"
/locus_tag="Pro0319"
/complement(3255..5378)
/locus_tag="chld"
/locus_tag="Pro0319"
/EC_number="4.99.1.1"
/codon_start=1
/transl_table=1
/product="Protoporphyrin IX Mg-chelatase subunit Chld"
/protein_id="AAP9365.1"
/db_xref="GI:33237297"
/translation="MWVPSKEKEPOSISIVLPSLEANTSPPLAAYVGHGTLKLIAMA
AVDPGLGVIIAGRGTKSYLARGHLALPPIEVLDRFNGDIQGRFPGRNIDPVE
PEWMDATNELIKORESNNSANGKODLQTRVOAPFVQVPLGTEDELVAVAVASL
DSGSAVFPQGLAAHARGVLYIDELNLDDGIIVMLAAVSGESGVERGLSLSHC
RPLIATYVPEGAARDHLDRFAIVLADQIISNEQREVEITSEALAFGSTSEPSK
MDEATQALATOLLAROMLSDIALSOGIEVLYKEAIRGVEGRSELVARYAKAA
ALCGRDKYARADIKYAVRIYVAPRSIQIPPEHMGPPDEOQPPPKRESIDEDIE
DEDIEDDMPEDQAPPIPEEFMDPEACADIDPULLPSAKSKSGSGSSGSAVFS
GRVYVKTLPKGPVTRIAVDAFLRAAPYQAREREPGRVIVQEGDLAKLQKKA
GSLVFIIVDAGSMALNMQSAKALIRLITEAVESRDEVSLLPFRGEQVLELLPFR
SITAKRLEITMPCCGGSPLAHLGLOARVGNALSTGDLQVVVVAITGGRGPIPG
KSLGPELIDGSDSVLDKOEVLIDATRYRTGIGIKLLVIDTERKFIASGIGDLADAAG
KYVOLPKATDOLASIMDAINTVK"
/complement(5387..6232)
/gene="tlg2c"
/locus_tag="Pro0320"
/complement(5387..6232)
/gene="tlg2c"
/locus_tag="Pro0320"
/codon_start=1

/transl_table=1
/product="ABC-type transport system involved in resistance
to organic solvents periplasmic component"
/protein_id="AAP9366.1"
/db_xref="GI:33237298"
/translation="MRRSFRPAIYGFSLIGSVVIFSGTLMLKGLKISNMTWYFANF
ADAGSLGCTVTRGQIOGVKTIITSCDVPARKISINNUNLVLPDYVAKIETSL
LGGDNEVMIQSGPILBISISNPKQENCPDLICADDSIKGNLDNLSLTGGINF
LNEENEBIKRIVNSIQEPDSTENDELIRSLKELFAPRPIDELKTVANTINI
LGSIDDPILSDISSSSISKSFPAKDKI SGKVDIEINDELINIGIKKAIIGKLF
NDIYE"
/complement(6238..7050)
/gene="tlg2a"
/locus_tag="Pro0321"
/complement(6238..7050)
/gene="tlg2a"
/locus_tag="Pro0321"
/codon_start=1
/transl_table=1
/product="ABC-type transport system involved in resistance
to organic solvents ATPase component"
/protein_id="AAP9367.1"
/db_xref="GI:33237299"
/translation="MKDLPSKPISSKRVYVENDSLMKWADNSLVNKVSIKMPGER
IATIGPSGCGSTVLRLLAGLLPSGDLKFLGMONYLRIDQTFPPDVRVLVQNPAL
IGSLVEENVGFLRKNSITDKRAREIVISCLEVGLYNIADKFPQLSGGMQKRV
FARALINDPKRSDMPLLFDEPTAGDPVACTRIEDLVKTTLLAMGCSIVVSHW
STIERSARVYMLVGEFQMGSDIDPFNTQNPPIIKQFRITSLGPMQKXI"
/complement(7227..8555)
/locus_tag="Pro0322"
/complement(7227..8555)
/locus_tag="Pro0322"
/codon_start=1
/transl_table=1
/product="uncharacterized conserved membrane-associated
protein"
/protein_id="AAP9368.1"
/db_xref="GI:33237300"
/translation="MLGRIQALGMLPGLVYKSMYLTSGLLIYALFGASIMADLRP
IYVVEGIFLLGITTKFLPSITGPVIFVIGISLLIWGQSRFSKSIONALAPKRDV
LIDALRAKSKNKGNGNIVAGIGGTGLSALLGLGRYSRIATVADADGSSGLRL
ELGVPPGDINDICLAALSTERSILTRLEQVPSGCGVGHSPGNILALTSITGNL
ETAITASSRVLAIVQGVVPATNADVRLMAELNENBRIGESALINVSPIRIGCYPE
QPPALPSAIEAIIENNELVIDGPGLYSLLENLVLPELVEALIQSKAKKLTICMLMT
PGETDGLDIAGHIRAIEAQLASFGVTRIRFNSIIAQKELPSALLEYLISGASPVIC
DNRRLRSQGYRLFKASLOESKIISSKSVLRHDPKRLSLAWMPYKRYKXA"
/complement(8556..9101)
/gene="ndhJ"
/locus_tag="Pro0323"
/complement(8556..9101)
/gene="ndhJ"
/locus_tag="Pro0323"
/codon_start=1
/transl_table=1
/product="NAD(P)-H-guinoine oxidoreductase subunit J"
/protein_id="AAP9369.1"
/db_xref="GI:33237301"
/translation="MNEBLDKKITEETNQEVLEPQRFISILLCKSGIEHVNEDHI
GIEIISVPEFLDIIIVSLKNGFNVLDCQGYEGPGLNTVCFYHLIATINYSKDSK
PREVAVKYLNDGDLKVPYVKIIFRGSDMQGQRETYDMFGVAFVDHPKRLMBEDW
KGMPLKXYIOPDPYEMODAY"
/complement(9104..9832)
/gene="ndhK"
/locus_tag="Pro0324"
/complement(9104..9832)
/gene="ndhK"
/locus_tag="Pro0324"

Query Match 28.9%; Score 566; DB 1; Length 300355;
Best Local Similarity 59.3%; Pred. No. 7.5e-154;
Matches 1004; Conservative 0; Mismatches 680; Indels 10; Gaps 2;
QY 139 GGGGCTTATATCTGATGATAGCCTGAACGCGCATGGGCTCAACACATTTTGGCTAT 198

Db 195975 GGGGAGATGCGCTTATGGAATCTCTTAAATCGACGGTGTGATACATATTTTGGGTAC 196034
 QY 199 CCGGCGGGGCAATTTTGGCCATCTATGATGTAAGTACCGCTTGAAGCGGGGGAA 258
 Db 196035 CCGAGAGAGCTATCTCTCTATGATGCGGTTTAAAGCAAGAAAAGCAGGTGG 196094
 QY 259 ATTGAGCATATTTGGTGGCCATGAACAAGAGCTTCCCATGCGCGAGTGGTATGCC 318
 Db 196095 GTAAACATATTTTGGTATGACATAGACAAGGTGGTCTCATGCGCATGATGATATGCA 196154
 QY 319 AGAGCACAGGTAAAGTGGAGTGTGTGGTATCATCTGAGCACGAGGCGCATACTG 378
 Db 196155 AGGTCCACAGGAAAGTGGTGTATGTTTGGACCTCAGACACGAGCAACAAATTTG 196214
 QY 379 GTGACCGGCAATGCGCAATGCCATTTGGACTCGGTGCCATGGTATTAAGAGAG 438
 Db 196215 GTTACAGGAATGCCACAGCTCAAAATGAGCTGTGTTCCATTGGTTATTACTGACAA 196274
 QY 439 GTGGGCGGTGCGCATGATGGTATGCGATGCTTCCAGAAATTTGACATTTTGGCATGCC 498
 Db 196275 GTTCCCAAGACAGCTATTTGGAATCTACGCTTTCAAGAAACAGATATTTTGGATTA 196334
 QY 499 TTACCGATCTTTAGACACTCTCTATGTTGATGAGTGGCGGAGATGGCTCGCATGGT 558
 Db 196335 CTCCCAATTTGTAACATTTCTGGGTGTAGAGATCTTCTGAATATGATCTGTATTT 196394
 QY 559 ACTGAGGCTTTCAATCTTGTCTAGCACCGGTGTCCTCGGCGCGGTGGTGTATGATTTCC 618
 Db 196395 GCGCAGGCTTTTATTTATGCTCTCTCTGCGGCCCAAGACCAAGTCTGTATGACATTTCA 196454
 QY 619 AAGATGTGGGCTTGAAGAAATGTAGTACATTTCCCTCGAACCCCGGAGAGTTAATCTA 678
 Db 196455 AAGATGTGGGCGCAAGAACTTTTATATATCACTGTTCAGCCAGGTTCTTTATCTCT 196514
 QY 679 CCGGCTTATGCGCCCAAGGTTAAAGTAAATCCCGCAAAATTAATGCGGATTTGCANTG 738
 Db 196515 GAGGCTTATCTCTCTCTCAAAAGCCAGACCGAGTCTATATCTAAGGCTGATTA 196574
 QY 739 TTGAGAGAGGCGAGAAATCCCTGTGCTACGTAAGGAGGAGGGCGATGGCCGCAATGCC 798
 Db 196575 ATTGAAAGAGCTTAAAGCCACTTTTATATGATGAGAGAGAGTATTTCTGCAAGTCT 196634
 QY 799 CATGCCAGGTGACAGAAATTTGCGAAAGGTTCCAGTTGCCGATACCAACCCCTGATG 858
 Db 196635 CATGACACTTATTGCTATTTGCCAAGGATATCATCTTGCTGTATGACAAACTGATG 196694
 QY 859 GGAATTTGGGGCTTTGAAGAAAACATCCCTTTGGTGGTATGTTGGTATGCAATGCC 918
 Db 196695 GGTAAAGGTGCAATTTGATGAACGTAATCTCTTCTGTGGGAGTGTAGGGATGCATGCC 196754
 QY 919 CACCGCTATGCCAATTTGCCGTACGCAATGTGATTTGTTGATTTGACAGTGGGGCCGT 978
 Db 196755 ACTGCATATGCCAATTTGCCGATACGTAATGCATTTATTTAGTATAGCATAGAGCAGT 196814
 QY 979 TTGACGACCGGTTACTGCGAACTAGACGAATTTGCTAGCCGCGCAAGTATTTAC 1038
 Db 196815 TTGACGACCGAGTACTGGAAGACTTGTATTTGACGCAAAATGCAAAAGTTGTAT 196874
 QY 1039 ATTGACATGACCCCGGAGGTGGGAAAAACAGGGCTCCGATGTGCCATTTGGGG 1098
 Db 196875 TTGTAAGTAGACCTGCGAAGATTAACAAAAACAGAGTTGCGAGTGTGGGTTTAGA 196934
 QY 1099 GATGTAGCCCATGTTTGAACAGCTTTGCAAGCGGGCCGCGGAATGGATTAACCCACC 1158
 Db 196935 GATTTAGCAATTAAGTCTTTGATTTACATGAATTAACAAAGAGA-----ACT 196985
 QY 1159 CATCCCATACGACCCGAGCATGGTTAATGCAATTTGATCTATGGCGGACGATTAACCC 1218
 Db 196986 GTTAAACCAATATCTCAAGAAATGGTGGCTCAAAATAGAGAGTGAAGATTTAAATATCA 197045
 QY 1219 CTCAGGTGCCCCACTATAGGATTAATGCCCCCCAGAGGTAGTACAGAAATGGT 1278

Db 197046 CTTTAAATACCTTCAAAAGAGGGGAATATATCAACAGAGATTTGATTAATTAGA 197105
 QY 1279 CGCAGGCCCCCGATGCTCTACTACCAACCATGATGTGGAGCAACACAAATGTGGGGGCC 1338
 Db 197106 GATTTAGCACAGATGCTTATATACATGATGTGGGCCAATCATTAATGTGGGACGT 197165
 QY 1339 CAGTTTGAACAATGGCCCCCGCCGATGATTTCCAGTGTGGCTTGGGTAGCATGGCC 1398
 Db 197166 CAATATTTAAGAAATAGTCCAGCTCATTTGATTAGACAGTGCAGATTTGGGAACAAGGGT 197225
 QY 1399 TTTGGTTTACCTGCGGCATAGGAGCAAAAGTGGAGTG- GGGAGAGAGCGGTCAATTTG 1457
 Db 197226 TTGGAATACCTGCTGCTTAGAGTGCAGATTTGCTTTGCCAAACAAACAGATTTTGC 197285
 QY 1458 ATCAGTGAAGATGCCAGCTTCCAAATGAATCTTCAAGAACTGGGAACCTTAGCCAGTAC 1517
 Db 197286 ATTGACAGGTGATGCAAGCATTTTGTATGAAATTTCAAGAAATTTAGGAGCTATCTGCAATAT 197345
 QY 1518 GACATCCAGTTAAACTATTTATCTCATTAACGTTGGCAGGGAGATGTGGCTCAGTGG 1577
 Db 197346 CAATTAATACAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 197405
 QY 1578 CAACAAATTTCTACAGAGACGTATTTCTGCTTTCAATGTCCTCAGGCGCATGCGAGAC 1637
 Db 197406 CAGCAAGTTTATATGATGAAGATTTACGCAACTATATGTTAAACGGGAGCCGAAT 197465
 QY 1638 ATTAATCTCTCTGTGAAGCTTATGSCATCAAGGTTATTAAGTGTGGCAGCGGAAAGAT 1697
 Db 197466 TTTGTGCTTTAGCTATAGTCTTTTGGAGTGGAGGGGTATTAATTAAGAGCGTAAATCA 197525
 QY 1698 TTGCGCCCGGCGATGCGCCGAATGCTAGCCCAAGTGTGCTGTGTAATGATGGTG 1757
 Db 197526 TTGAAAGAAAGTTAAAGAGCTTTAATTCGAAAGACCAATGCTCATTTGATTTCA 197585
 QY 1758 GTCAAAAAGATGAAATGTTTACCTATGATTTGCCCGGACATGATGATGCGCAATG 1817
 Db 197586 GTCCGTAGAGATGAGATTTGTATCAATGTTCCACAGGTAAAGCAATGCTGAATG 197645
 QY 1818 CTAGTTTACCGGA 1831
 Db 197646 GTTGGTTGCTTA 197659

Search completed: July 25, 2004, 10:17:16
 Job time : 7875.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 03:57:37 ; Search time 822.718 Seconds

(without alignments)
10115.521 Million cell updates/sec

Title: US-09-893-033-6

Perfect score: 1959

Sequence: 1 gccatagagagccatgcgcg.....cataagccaaattgaattc 1959

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Geneseq_29Jan04:*
2: Geneseqn1980s:*
3: Geneseqn1990s:*
4: Geneseqn2000s:*
5: Geneseqn2001as:*
6: Geneseqn2002as:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1958	99.9	1959	6	AAD27607 Synchoc
2	454.6	23.2	1728	6	ABQ90234 M. capsul
3	415	21.2	1728	6	ACA53857 Prokaryot
4	382.6	19.5	1755	7	ACF71672 Photorhab
5	382.6	19.5	110000	7	Continuation (49 o
6	382.6	19.5	110000	7	Act65387 Photorhab
7	381.6	19.5	1878	7	ACA39863 Prokaryot
8	379.6	19.4	1722	7	ACA37690 Prokaryot
9	379.6	19.4	1722	7	ACA53435 Prokaryot
10	378.8	19.3	1875	7	ACA23942 Prokaryot
11	373.8	19.1	1725	7	AAS54310 Pseudomon
12	373.8	19.1	1725	7	ACA42854 Prokaryot
13	368.6	18.8	1734	2	AAT45893 R. capsul
14	368.6	18.8	1815	4	AAS52288 E. coli D
15	368.6	18.8	1815	7	ACA32352 Prokaryot
16	368.2	18.8	1833	7	ACA39024 Prokaryot
17	368.2	18.8	96109	4	AAZ28548 Genomic f
18	367.8	18.8	1725	7	ACA43064 Prokaryot
19	366.8	18.7	1854	7	ACA38394 Prokaryot
20	366.8	18.7	1857	7	ACA40769 Prokaryot
21	365.8	18.7	110000	4	AA199682_33 o
22	365.4	18.7	110000	4	Continuation (34 o
23	363	18.5	1596	7	ACA44223 Prokaryot

24	363	18.5	1725	6	AAS97210	AAS97210 Neisseria
25	363	18.5	1725	6	AAS97243	AAS97243 Neisseria
26	363	18.5	1728	7	ACA41957	ACA41957 Prokaryot
27	358.2	18.3	13732	3	AAA81484	AAA81484 N. mening
28	358.2	18.3	110000	3	AAA81489_1	Continuation (2 of
29	358.2	18.3	110000	3	AAA81489_2	Continuation (3 of
30	358.2	18.3	349980	3	AAZ1611	AAZ1611 Neisseria
31	355.2	18.1	110000	6	ABA03041_20	Continuation (21 o
32	355	18.1	1725	7	ABZ38433	ABZ38433 N. gonorr
33	355	18.1	1725	7	ACA41281	ACA41281 Prokaryot
34	354	18.1	1722	7	ACA36592	ACA36592 Prokaryot
35	350.8	17.9	6506	6	ABQ71032	ABQ71032 listeria
36	349.8	17.9	1761	7	ACA25532	ACA25532 Prokaryot
37	345.8	17.7	1554	7	ACA45757	ACA45757 Prokaryot
38	345	17.6	2286	6	ABQ69168	ABQ69168 listeria
39	345	17.6	110000	6	ABQ69245_20	Continuation (21 o
40	345	17.6	110000	6	ABQ69245_21	Continuation (22 o
41	345	17.6	319630	6	ABQ67194	ABQ67194 listeria
42	342.2	17.5	1713	7	ACA27233	ACA27233 Prokaryot
43	338.8	17.3	1725	4	AAS00992	AAS00992 B. subtil
44	338.8	17.3	3291	4	AAS02313	AAS02313 B. subtil
45	337.8	17.2	1956	7	ACA29658	ACA29658 Prokaryot

ALIGNMENTS

RESULT 1
AAD27607
ID AAD27607 standard; DNA; 1959 BP.
XX
AC AAD27607;
XX
DT 18-APR-2002 (first entry)
XX
DE Synchocystis sp. AHA5 large subunit gene, 11vg.
XX
KW Cyanobacteria; acetylhydroxyacid synthase; AHA5; acetoacetate synthase;
KW phytoene desaturase; PDS; plant trait control; plasmome transformation;
KW herbicide resistance; fungicide resistance; insecticide resistance;
KW plant growth protectant; ds.
XX
OS Synchocystis sp.
XX
PN WO200200915-A2.
XX
PD 03-JAN-2002.
XX
PF 27-JUN-2001; 2001WO-US020338.
XX
PR 27-JUN-2000; 2000US-0214705P.
XX
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Kakefuda G, Koop H, Sturmer S, Zhen R;
XX
XX WPI; 2002-139930/18.
DR
XX
XX
PT New cyanobacterial nucleic acid fragments encoding acetylhydroxyacid
PT synthase (AHA5) or phytoene desaturase (PDS); useful for conferring
PT herbicide, fungicide or insecticide resistance, and for identifying AHA5
PT or PDS inhibitors.
PS
PS Claim 1; Page 35-37; 70pp; English.
XX
XX The invention relates to cyanobacterial nucleic acid fragments encoding
XX acetoacetate (acetylhydroxyacid) synthase (AHA5) large and small subunits
XX or mutant phytoene desaturase (PDS). Polynucleotides of the invention are
XX useful for controlling plant traits via nuclear or plasmome
XX transformation, e.g. for conferring herbicide, fungicide or insecticide
XX resistance. Particularly, they are useful in identifying novel PDS and
XX AHA5 inhibitors and in plant transformations for conferring resistance
XX and cross-resistance to certain bleaching herbicides and AHA5-inhibiting

CC herbicides. The present sequence is *Synechocystis* sp. strain PCC 6803
 CC AHAS large subunit gene, *llvG*
 XX
 SQ Sequence 1959 BP, 465 A; 493 C; 531 G; 469 T; 0 U; 1 Other;
 Query Match 99.9%; Score 1958; DB 6; Length 1959;
 Best local Similarity 100.0%; Pred. No. 0;
 Matches 1959; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCATGAGAGCCCATCCCGCATTTGAGTTCAATTGAAAGCATTTAGGCTTACGCTTCTTA 60
 DB 1 GCCATGAGAGCCCATCCCGCATTTGAGTTCAATTGAAAGCATTTAGGCTTACGCTTCTTA 60
 QY 61 ACCGATTGTCAGTGGTTGTCATCAATTCCTAATCCCAAAACAAATTTCTGAAAACGTGT 120
 DB 61 ACCGATTGTCAGTGGTTGTCATCAATTCCTAATCCCAAAACAAATTTCTGAAAACGTGT 120
 QY 121 CCTAGCCAAAGCGAAACCGGAGCTTATATCTGATGATAGCTGAAAGCCATGGGCTC 180
 DB 121 CCTAGCCAAAGCGAAACCGGAGCTTATATCTGATGATAGCTGAAAGCCATGGGCTC 180
 QY 181 AAACACATTTTGGCTATCCCGGCGGGCAATTTGGCCATCTATGATGAACGTACCGC 240
 DB 181 AAACACATTTTGGCTATCCCGGCGGGCAATTTGGCCATCTATGATGAACGTACCGC 240
 QY 241 TTTGAGCGCGCGGGGAAATTTGACATATTTGGTGGCCATGAAACAGAGCTTCCCAT 300
 DB 241 TTTGAGCGCGCGGGGAAATTTGACATATTTGGTGGCCATGAAACAGAGCTTCCCAT 300
 QY 301 GCGCGGATGGGTATGCGAGAGCCACAGGTAAAGTGGAGTTGTTGGTACATCTGA 360
 DB 301 GCGCGGATGGGTATGCGAGAGCCACAGGTAAAGTGGAGTTGTTGGTACATCTGA 360
 QY 361 CCAGGGGCGACTAATCTTGATGACCGGATTTGCCAATGCCATTTGGACTTGGTCCCATG 420
 DB 361 CCAGGGGCGACTAATCTTGATGACCGGATTTGCCAATGCCATTTGGACTTGGTCCCATG 420
 QY 421 GTGGATTTACTGAGAGGTGGCGCGGTCATGATGTTGAGCAATGCTTTCAGGAAAT 480
 DB 421 GTGGATTTACTGAGAGGTGGCGCGGTCATGATGTTGAGCAATGCTTTCAGGAAAT 480
 QY 481 GACATTTTGGCATCACTTACCGATCGTTAAGCACTCTATGATGATGATGATGCGGCG 540
 DB 481 GACATTTTGGCATCACTTACCGATCGTTAAGCACTCTATGATGATGATGATGCGGCG 540
 QY 541 GATATGCTGGCATTTGTTACTGAGGCTTTCATCTTGTCAAGCAACGCTGCTCCGCGCG 600
 DB 541 GATATGCTGGCATTTGTTACTGAGGCTTTCATCTTGTCAAGCAACGCTGCTCCGCGCG 600
 QY 601 GTTTGATGATATTTCCCAAGATGTGGGCTTAGAAGATGTGATACATTCCTCCGAC 660
 DB 601 GTTTGATGATATTTCCCAAGATGTGGGCTTAGAAGATGTGATACATTCCTCCGAC 660
 QY 661 CCGGATGACTTAATCTACCGGGTTATGCGCCACGCTTAAAGTAAATCCCGACAAAT 720
 DB 661 CCGGATGACTTAATCTACCGGGTTATGCGCCACGCTTAAAGTAAATCCCGACAAAT 720
 QY 721 AATCGGATGCAATGTTGAGAGGCGCAAAATCCCTGCTCAACGATGGGAGGAGG 780
 DB 721 AATCGGATGCAATGTTGAGAGGCGCAAAATCCCTGCTCAACGATGGGAGGAGG 780
 QY 781 GCGATGCGCCCAATGCGCAATGCGCCAGAGTGCAGAAATTTGCGGAAAGGTTCAGTGGC 840
 DB 781 GCGATGCGCCCAATGCGCAATGCGCCAGAGTGCAGAAATTTGCGGAAAGGTTCAGTGGC 840
 QY 841 GTAAACAACCACTGATGGGAAATTTGGGCTTTTGAAGAAACCATCCCTTCGAGGAGT 900
 DB 841 GTAAACAACCACTGATGGGAAATTTGGGCTTTTGAAGAAACCATCCCTTCGAGGAGT 900
 QY 901 ATGTGGGATGATGCGCAACCGCTATGCGCACTTTGCGGTGACGAAATGTGATTTGTTG 960
 DB 901 ATGTGGGATGATGCGCAACCGCTATGCGCACTTTGCGGTGACGAAATGTGATTTGTTG 960

QY 961 ATTGAGTGGGGGCGCGTTTGGACAGCGGGTAACTGGCAAACTAGACGAATTTGCTAGC 1020
 DB 961 ATTGAGTGGGGGCGCGTTTGGACAGCGGGTAACTGGCAAACTAGACGAATTTGCTAGC 1020
 QY 1021 CCGCCAAAGTAAATTCATTTGACATGACATCCCGGAGAGTGGGAAAAAACAAGGCTCC 1080
 DB 1021 CCGCCAAAGTAAATTCATTTGACATGACATCCCGGAGAGTGGGAAAAAACAAGGCTCC 1080
 QY 1081 GATGAGCCATGTTGGGGGAGTAAAGATGATGATTTTGAACAGCTTTTGAAGGCGGCGG 1140
 DB 1081 GATGAGCCATGTTGGGGGAGTAAAGATGATGATTTTGAACAGCTTTTGAAGGCGGCGG 1140
 QY 1141 GAATGATTAACCCCAACCATCCCATACCAAGCCAGCATGTTAAATGCAATGATCAT 1200
 DB 1141 GAATGATTAACCCCAACCATCCCATACCAAGCCAGCATGTTAAATGCAATGATCAT 1200
 QY 1201 TGGCGGACCGATTAACCCCTCTCAGGTGCCCCCACTATGAGATTAATTTGCCCCCAAG 1260
 DB 1201 TGGCGGACCGATTAACCCCTCTCAGGTGCCCCCACTATGAGATTAATTTGCCCCCAAG 1260
 QY 1261 GTAGTACACGAATTTGTCGCGCAGGCGCCCGCATGCTACTACACCGATGTGGACAA 1320
 DB 1261 GTAGTACACGAATTTGTCGCGCAGGCGCCCGCATGCTACTACACCGATGTGGACAA 1320
 QY 1321 CACCAATGTGGGCGGCGCCAGTTTGAACATGCGCCCGCATGATTTCCAGTGTCT 1380
 DB 1321 CACCAATGTGGGCGGCGCCAGTTTGAACATGCGCCCGCATGATTTCCAGTGTCT 1380
 QY 1381 GCGTTGGGTAAGATGGGCTTTGTTGTTTACCTGCGCCATGGGAGCCAAAGTGGAGTGG 1440
 DB 1381 GCGTTGGGTAAGATGGGCTTTGTTGTTTACCTGCGCCATGGGAGCCAAAGTGGAGTGG 1440
 QY 1441 GACGAGCGGTCAATTTGATCAGTGAAGATGCGACCTTCCAAATGATCTTCAGAACTGG 1500
 DB 1441 GACGAGCGGTCAATTTGATCAGTGAAGATGCGACCTTCCAAATGATCTTCAGAACTGG 1500
 QY 1501 GAACCCCTGCCCCAGTAAAGCAATCCAGGTTAAATTAATTTTCCAAATGACGTTGGAGG 1560
 DB 1501 GAACCCCTGCCCCAGTAAAGCAATCCAGGTTAAATTAATTTTCCAAATGACGTTGGAGG 1560
 QY 1561 GGAATGTCGTCAGTGGCAAACTTCTCTGTAAGGCTTATGCTTCTTAACATGT 1620
 DB 1561 GGAATGTCGTCAGTGGCAAACTTCTCTGTAAGGCTTATGCTTCTTAACATGT 1620
 QY 1621 CCGAGGATGCGCAACATTAATCTCTCTGTAAGGCTTATGCTTCTTAACATGT 1680
 DB 1621 CCGAGGATGCGCAACATTAATCTCTCTGTAAGGCTTATGCTTCTTAACATGT 1680
 QY 1681 TGGCAACCGGGAAGATTTGGCGCGGATCGCGCAATGCTTACCCCAATGCTGCTG 1740
 DB 1681 TGGCAACCGGGAAGATTTGGCGCGGATCGCGCAATGCTTACCCCAATGCTGCTG 1740
 QY 1741 TGGTATGATGTGTGTGTAAGAAAGTGAATCTTTAACCCTATGATTTGCCCGGCA 1800
 DB 1741 TGGTATGATGTGTGTGTAAGAAAGTGAATCTTTAACCCTATGATTTGCCCGGCA 1800
 QY 1801 TGAATGATGATGATGTGTGTAAGAAAGTGAATCTTTAACCCTATGATTTGCCCGGCA 1860
 DB 1801 TGAATGATGATGATGTGTGTAAGAAAGTGAATCTTTAACCCTATGATTTGCCCGGCA 1860
 QY 1861 TGGTGAAGTGAACCAATTTGCCAAACCCCAAAATTTCAATCAACCAATGCTTGTCTG 1920
 DB 1861 TGGTGAAGTGAACCAATTTGCCAAACCCCAAAATTTCAATCAACCAATGCTTGTCTG 1920
 QY 1921 GTGAGGCCAAATCTTAACCCCAATTAAGCCAAATTTGAATTC 1980
 DB 1921 GTGAGGCCAAATCTTAACCCCAATTAAGCCAAATTTGAATTC 1980

RESULT 2
 ABO90234
 ID ABO90234 standard; DNA; 1728 BP.
 XX

AC AB090234;
XX
DT 01-OCT-2002 (first entry)
XX
DE M. capsulatus gene #219 for DNA array.
XX
KM Micro array; gene; ds; differential expression; gene expression.
XX
OS Methylococcus capsulatus.
XX
PN W0200255655-A2.
XX
PD 18-JUL-2002.
XX
PF 14-JAN-2002; 2002WO-N0000019.
XX
PR 12-JAN-2001; 2001NO-00000235.
XX
PR 12-JAN-2001; 2001NO-00000239.
XX
PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
XX
PI (TIGR-) TIGR.
XX
PI Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;
PI Salzberg SL;
XX
DR WPI; 2002-557818/59.
XX
XX
XX Novel DNA array useful for determining differential expression of
PT Methylococcus capsulatus genes, comprises polynucleotides or
PT oligonucleotides representative for a selective number of Methylococcus
PI capsulatus genes.
XX
XX
XX Claim 19; Page 156-157; 678BP; English.
XX
XX
XX The invention relates to a novel DNA array giving a representation of a
CC number of Methylococcus capsulatus genes. The method of the invention is
CC useful for determination of the differential expression of the genes of
CC M. capsulatus, and for studying gene expression on a genomic scale and in
CC gene expression assays of M. capsulatus genes. The sequences shown in
CC AB090016-AB091855 represent M. capsulatus genes for use in arrays of the
CC invention
XX
SQ Sequence 1728 BP; 356 A; 576 C; 506 G; 290 T; 0 U; 0 Other;

Query Match 23.2%; Score 454.6; DB 6; Length 1728;
Best Local Similarity 56.8%; Pred.No. 2.6e-133;
Matches 962; Conservative 0; Mismatches 709; Indels 22; Gaps 6;

QY 136 ACCGGGCTTATATCTGATGATAGCTGAAACCGCATGGGGTCAACACATTTTGGC 195
DB 10 AGCGGTGGGAAATGCTGTCATATGCTTAAAGACGAGGGTGTGATATCATCTTGGGC 69
QY 136 TATCCCGGGGGCAATTTTGGCCATCTATGATGAAGTACCGCTTTGAAGCGGCGG 255
DB 70 TATCCGGGGGCTGAGTACTGATATTAAGATGGCTGTTCAAACAGAGAGATG----- 124
QY 256 GAAATTGAGCATATTTGGTGGCGCATGAACAAGAGCTTCCCATGGCGGATGGGTAT 315
DB 125 ----TAAAGCATCTCTGTTGGCGCAACAGAGGCGGACCCGCGGCGGATAC 180
QY 316 GCCAGAGCCACAGTAAAGTGGAGTTTGTTCGATCATCTGACACAGGCGGCACTAAC 375
DB 181 GCCCCTCCACCGGCAAAACCCGCGTGTCTGTGACTTGGGGCCCGGTGCGACCAAT 240
QY 376 TTGTTGACCGGATGCGCAATGCGATTGACCTGGTGGCCATGGTGGTGAATTATGGA 435
DB 241 GCGGTACCGGATGCGCAACCGCTTACATGATTTGATTCGCGTGTGTGATCAACGAGC 300
QY 436 GAGGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
DB 301 CAGGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 496 ACCTTACGATCGTTAAGACCTCTATGTGTGATGATGCGGCGGATATGCGTCGAT 555
DB 361 ACTGCGCGGCGGTGAAGCAAACTTCTGTGTCAAGACATCCAAAGCTGGCCGAACT 420
QY 556 GTTACTGAGCTTTTCCATCTTGTAGACCGGTCGTCGCGGCGGCTTTGATGATGAT 615
DB 421 TTCAAGAGGCTTCTTACATCGGACACGAGGAGCGCGCGCGCTGTGTGTGATGATG 480
QY 616 CCCAAGATGTGGGCTTGAAGATGTGATCATTTCCCTCGACCCCGGTGATGAT 675
DB 481 CCCAAGATGTGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCG 540
QY 676 CTACCGGATTTATGCGCCACCGTTAAAGTATCCCGCAAAATTAATGCGGATTTGCA 735
DB 541 CTGAGGTCTTCAATCTTCTGTTGAAGGCGCATCCATCCATCAATCAAGAGGCGGTGAG 600
QY 736 TTGTTGAGAGGCGCAAAATCCCTGTCTCTAGTGTGAGGAGGAGGCGATGCGCGCAT 795
DB 601 CTATGCTGTGCGCTGCGCGCGCGCGATGATCTACAGCGGCGCGCGCTATCTCGGCAT 660
QY 796 GCCCATGCGCAGGTGAGAGATTTGCGGAAAGTTCCAGTTGCCGTAAACAACCTTG 855
DB 661 GCCCGGACGAACTGACCAAGCTGACGCGGCTGTGATCTCCCATCAACCAACCTTG 720
QY 856 ATGGGAATTTGGGCGCTTTTGAAGAAACATCCCTTTGGTGGGATGTTGGGTATGAT 915
DB 721 ATGGGTCTGGGCGCTTATCCCGCACGACCGCATGCTGTGCGATGCTGGGATGAC 780
QY 916 GGCACCGCTATCCCACTTTGGCGTCAAGCATGATGATTTGTGATGCAAGTGGGCGC 975
DB 781 GGCACCTAGAGCGCAACATGCGATGACGATGCGAGCTGTGCTGCGCGCGCGCGC 840
QY 976 CGTTTGAAGACCGGCTTATGCGCAACTGACAGATTTGCTAGCGCGCGCAAGTAT 1035
DB 841 CGCTTGAATGACCGGCTTATGCGCAACTGACAGATTTGCTAGCGCGCGCAAGTAT 900
QY 1036 CACATTGACATGACCGCGCGGAGTGGGAAAAAACAAGGAGTCCCGATGCGCATTTGT 1095
DB 901 CACATGACGATGACCGCGCGCTGATCTCCAAACCGTGAAGTGAACATCCCATCTGTC 960
QY 1096 GGGGATGTACGCAATGTTTGAACAGCTTTTGAAGCGCTTTGACGCGGCGCGGAA 1155
DB 961 GCGGAGTGTGCGCGCGGCTGTGCGGACATGATGACATGATCAAGCTTCCGCGCGC 1020
QY 1156 ACCATCCCATACCAACCGGATGATTAATGCAATGATGATGCGGCGGACGATTC 1215
DB 1021 CCGATGCGGAGCGCGCTGAAAGCGCTGTGGCAAGATGACATGCGGCGGCTGCGAC 1080
QY 1216 CCGCTCCAGGTGCGCACTATGAGATTAATGCGCGCGGAGGAGTATGACGAAAT 1275
DB 1081 TGCTTGGCTACGACCGGACGACGACGAGATCAACCGCAGTTCTGTATCGAACATTA 1140
QY 1276 GTTGGCGAGGCGC---CGATGCTCTACACCAAGGATGTGGGACAAACCAATGTG 1332
DB 1141 TGGAACTGACGCGGCGGACGCTTGTGCGCTTCCGAGCTGTGTGACACAGATGTG 1200
QY 1333 GCGGCGCAGTTTGA---CAATGCGCGCGCGCGGATGATTTTCCAGTGTGCTTGGT 1389
DB 1201 GCGGCGCAATACACAAATTCACAGAGCGCGCGCTGTGGTCAATTCGCGCGCTGGG 1260
QY 1390 ACCATGCGCTTGTGTTTACTGCGCGCATGAGACCAAGTGTG-AGTGGGAGACGAGG 1448
DB 1261 ACATGCGGCTTCCGATCTGCGCGCGGATCCGATCAATTTGCGGATCCGAGGAAGG 1320
QY 1449 GTCAATTGATCAGTGAAGATGCGACCTTCAATCAATCTTCAAGAACTGAGAACCTTA 1508
DB 1321 GTATCTGATTCACCGGCGAGGAGGACATCCAAATGATCCAGAACTTGGCACTGCGC 1380
QY 1509 GCCCATGACGATCCAGGTTAAACTATTAATTCATTAACGTTGGGAGGAGTGTG 1568
DB 1381 TTGCAGTACCGGACCGCGGCAAGATGCTCAACCTTAACACGCTTACATGGGCAATGG 1440
QY 1569 CGTACGTCGCAAAACCTTTCTACGAAGAACGTTATTTCTGCTTCTAATCATGTCCAGG 1628

DB 1441 CGACATATGGCAGGATTCATCATATGAAGCCGCTATTCGATCTCTACTTG---GAGACT 1497
QY 1629 ATGCCAGACATTAATCTCTCTGTGAAGCCATATGAGATCAAGGATATTAATCTGTGCGGAG 1688
DB 1498 CTGCGGACTTCTGCAAACTGGCGAGGCTTACGTCACGTCCGATGCCATGACAGCA 1557
QY 1689 CGGGAAGATTTGGCCCGCGCATGCCGGAATGTATGCCCAATG---GTCTGTGTGT 1745
DB 1558 CCGGCGGAGCTGCGCGCCGCTGGAAGAACCTGGAACCTCAAGACCGACCGGTGTTC 1617
QY 1746 ATGATGTGTGTGTCAAAAAGATGAAACCTGTATACCTATGATGCCCCCGGATAGT 1805
DB 1618 ATGATTTCTTCCACCGATCCCAACGGAAGAGTACCCCATGTGATCCGGAAGGCC 1677
QY 1806 AATGCCCAATGC 1818
DB 1678 CATCACGAGATGC 1690

RESULT 3
ACAS3857
ID ACAS3857 standard; DNA; 1728 BP.
XX AC ACAS3857;
XX
XX 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #35514.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Versinia pestis.
XX
XX WO20027183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948933.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX
XX P-PSDB; ABU43987.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
PS
PS Claim 14; SEQ ID NO 41727; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a gene
CC product's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 1728 BP; 416 A; 387 C; 473 G; 452 T; 0 U; 0 Other;
Query Match 21.2%; Score 415; DB 7; Length 1728;
Best Local Similarity 55.8%; Pred. No. 1,1e-120;
Matches 929; Conservative 0; Mismatches 705; Indels 31; Gaps 6;
QY 172 CATGGGCTCAACACATTTTGGCTATCCCGCGGCGCAATTTGCCATCTATGATGA 231
DB 49 CAGGGCGTTAAGACATGATTCGGTTATCCGCGGCGCGCACTGATATTTAAGATGCC 108
QY 232 CTGTACCGCTTTGAAGCGCGGCGGAATGAGCATATTTTGGCGGCATGAACAAGA 291
DB 109 CTGCAC-----ACGGTGGAGGCACTGATACGTCGTGTGGCGCAAGAACAGT 159
QY 292 GCTTCCCATGCGCGGATGGATGATGCCAGACCAAGGTAAAGTGGAGTTTGTTCGT 351
DB 160 GCGGTTCAATGCGCGATGATGATGCGCGGCAACGCGCGAGTGGTGTGATGCTC 219
QY 352 ACATCTGACACAGGCGGCGCATTAATCTGTGACCGGCAATGACATGCGCATTTGAGCTGC 411
DB 220 AGCTGTGCTCTGTGCAACCAATGCGATTCACCGTATTTGACCGGCTATATGATTA 279
QY 412 GTGCCCATGTGTGATTAATCTGAGAGGTGGGCGGTGCATGATTTGGATGCGATCTTC 471
DB 280 GTGCCAATGTGTGATCTTTGAGGCAAGTGCACAGCTCATGATTTGTTGATGCGCTC 339
QY 472 CAGGAATATGACATTTTGGGATCACTTACCGATCGTTAGACACTCTATGTGTAGT 531
DB 340 CAAGAATGTGACATGTGGGATTCCTCCGCGAGTGTCAACATAGCTTCTTGTTAAG 399
QY 532 AGTGGCGGAGATGAGTCTCGCATTTGTTACTGAGGCTTTCATCTGTAGACCGGTCGT 591
DB 400 CGCAACAGAAACATCCGATGATGTGTAAGAAAGCTTTTATCTGGGATCCACAGCGCGT 459
QY 592 CCGGCGCGGTTTGTATGATTAATTCCTCAAGATGTGGCTTGAAGAAATGATGATCTT 651
DB 460 CCGGCTCGGTGTATGATGACTTGGCAAAACATTTGCGGCCACGAGATTAATAGCT 519
QY 652 CCGCTGACCGCGGAGATTAATCTACCGGTTATGCGCCACGATTAAGATATGCC 711
DB 520 TAGGCTTATCCAGAGCAAGTCAATTTGGCTCTCTATATCTTAGGTGCAAGGCAACCT 579
QY 712 CGACAAATTAATGCGGATTCGCAATTTGTTGAGAGCGCCAGAAATCCCTTGTCTAGTA 771
DB 580 GGGCAGATTAATGCGGATTCGCAATTTGTTGAGAGCGCCAGAAATCCCTTGTCTAGTA 771
QY 772 GGGGAGGGGGAGATGCGGCAATGCGCATGCGGAGGAGGAGATTTGGGAAAGGTTG 831
DB 640 GATGTGTGTGATTAATGCGGCTTGTGCAATTAAGATTTAGATTTGCGAGAAAGTGT 699
QY 832 CAGTTGCGGATTAACACCACTGATGGAATTTGGGCTTTTGAAGAAACATCCCTT 891
DB 700 AACTTACCGATTAACAGAGCTCTGTATGAGGCTTAGGCAAGTTCCCTGTAGACATGCGCAA 759

[illegible]

RESULT	4
ACF71672	
ID	ACF71672 standard; DNA; 1755 BP
XX	
AC	ACF71672;
XX	
DT	20-NOV-2003 (first entry)

XX Photorhabdus luminescens nucleotide sequence #10139.
XX
DE Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
XX detection; food; gene expression; plant; animal; microorganism; toxin;
KM antibiotic; biopesticide; virulence factor; disease model; plague;
KM whooping cough; gene; ds.
XX
XX Photorhabdus luminescens.
OS
XX WO200294867-A2.
XX
XX 28-NOV-2002.
XX
XX 07-FEB-2002; 2002WO-IB003040.
XX
XX 07-FEB-2001; 2001FR-00001659.
XX
XX (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
XX Duchaud E, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A,
PI Buchrieser C;
XX
XX WPI: 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
XX
XX Claim 2, SEQ ID NO 10139; 1205bp; French.
XX
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
XX
XX Sequence 1755 BP; 484 A; 322 C; 459 G; 490 T; 0 U; 0 Other;
SQ
XX
XX Query Match 19.5%; Score 382.6; DB 7; Length 1755;
XX Best Local Similarity 54.6%; Pred. 2.2e-110;
XX Matches 877; Conservative 0; Mismatches 709; Indels 19; Gaps 5
XX
XX
XX 121 CTTAGCCAAAGCGCAAAACCGGGCTTATATCTGTATGATACCTGAAAGCCATGGGCTC 180
XX
XX Db CCGATGGAAGATGTATTCAGAGACTGAAATGGTGTCTCGATGTTATATGATCGAGGGT 84
XX
XX 181 AAACACATTTTGGCTATCCCGCGGGGCAATTTTGGCCATCTATGATGAACGTATCCG 240
XX
XX Db AACCATATATTCGTATTATCCGGGGGGGCACTAGCATGATTTATATGATGCCCTGAT--- 141
XX
XX 241 TTGGAAGCGGGGGGAAATTGAGATATTTTGGTGGCCATGAAACAGAGCTTCCCAT 300
XX
XX Db 142 -----ACGAGCTGGGGGCAATTGAGATATCTCGTAGTCATGAGCGGGGGCAATTCA 195
XX
XX 301 GCGCGGATAGGGTATCCAGAGCCACAGATGAAGTGGAGATTGTTCGGTACATCTGGA 360

Db 196 ATGCTGATGTTATGCTCGATTAACCGGGAAGTTGTTGATTTATACCTACGT 255
Qy 361 CCAGGGGCGACTAATTGGTGAACCGGATTTGCCAATGCCATTGGATCGTGGCCCAT 420
Db 256 CCAGAGAACCAATGCGCATTAACCGGATTTGCCAATGCCATTGGATCGTGGCCCAT 315
Qy 421 GTGGTATTTACTGAGAGGAGGCGCGGCGCATGATTTGGAGCGATTTCCAGGAATT 480
Db 316 GTTGTATTTGCTCGGTCAGTCCAGATTCGTTATGCGTATGATGATTCAGGAATT 375
Qy 481 GACATTTTGGCATCACCTTAACCGATCGTTAAGCACTCTATGTTGATGATGCGGCG 540
Db 376 GACATGTAAGGATATCCCGCCGATCGTAAACAGTTTATTTAGTGAAGAAAGAGAG 435
Qy 541 GATATGCTGCGCATTTGTTACTGAGCTTTCCATTTGCTAGCAACCGGTGTCGCGGCG 600
Db 436 GATATTCGGAAGGACTGAAAAAGCTTTCTATCTACTCTTCAGTGTCTGCGGAGACC 495
Qy 601 GTTTGATGATATTTCCCAAGATGATGAGGCTTAGAAGATGATGATCATTTCCCTGAC 660
Db 496 GTGCTATTTGATTTACCAAAAGACAGGTAATCCGCACTAATAATACCTATTTAT 555
Qy 661 CCCGTAAGCTTAATCTACCGGATTAATGCCGCCACCGTTAAAGTAAATCCCGACAAAT 720
Db 556 CCTGATCAGATTAAGTATGCGTTCTTATATCAACGATTCAGAGACCAAGGCGAGATT 615
Qy 721 AATCGGATTTGCAATTTGTTGAGACGAGGCAAAATCCCTTGTCTCTACGAGGAGGAG 780
Db 616 AAGCGGCGATTAAAGAACTATTAATGCGAAAAACAGTCAATATATGATGCGGCGG 675
Qy 781 GCGATCCCGCAATGCGGATCCCGAGGTCGAGAAATTTGCGGAAGTTCCAGTTGCCG 840
Db 676 GCGATTAACGAGATTTGTCAGAGAAATTAAGTATTTGCGGAAGAAATCAATATCCCT 735
Qy 841 GTAAACAACCAACCGATGAGGAAATTTGAGGAGGCTTTGACAGAAACCATCCCTTGGTGGG 900
Db 736 GTTGTGAGCTTCCGTAAGGCTTAGGCAATTTCCCTGCACTCATCCGCAACCTGGCG 795
Qy 901 ATGTTGGTATGATGAGCCACCGCTATGCAATTTGCGGTGAGCGAATGATTTGTTG 960
Db 796 ATGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
Qy 961 ATTGCAATGAGGCGCGCTTTGAGACGCGGTAATGCGCAATGAGCAATTTGCTAGC 1020
Db 856 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915
Qy 1021 CGCGCCAAAGTATTTCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 916 GAACCAACGATTTATCAATTTGATGATGATGATGATGATGATGATGATGATGATG 975
Qy 1081 GATGTCCTCATTTGTTGAGGAGTATGAGGCAATGTTTGAACAAGCTTTTGACGCGGCGCG 1140
Db 976 GATGATCCTATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1035
Qy 1141 GAATTTGATTTCCCAACCGATTTGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1036 ACAATGAAGACACACAGATTCAGATGCGTAAAGATTTGGTGTATTTATTCAGCAA 1095
Qy 1201 TGGCGAGCCGATTAATGCTTATGATTTGATGATGATGATGATGATGATGATGATG 1260
Db 1096 TGGCGAGCCGATTAATGCTTATGATTTGATGATGATGATGATGATGATGATGATG 1155
Qy 1261 GTAGTACACGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
Db 1156 GCTTTCAGGTTCTCTATGCTTACAGAGGAGAGACTTATGATCTTCAATGTTGGG 1215
Qy 1318 CAACACCAATGTTGAGGCGCGCATTTTGAACAATGG---CCCCCGCGATGATTTCC 1374
Db 1216 CAGCATCAGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
Qy 1375 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1433
Db 1276 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335

Qy 1434 ACTGGGGAGACAGCCGCTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1493
Db 1336 AAACCGAGAGCAGCGTATGCTGTTACGGGGAGTGTATGATGATGATGATGATGAT 1395
Qy 1494 GAATGGAACCCCTAGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
Db 1396 GAGCTGTCCACCGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
Qy 1554 TGGCAGGAGATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1613
Db 1456 TTCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1515
Qy 1614 AACATGTCACCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1673
Db 1516 TATATG---GATCTTCTGCTGATTTTGTCAAAATGCTGATGATGATGATGATGAT 1572
Qy 1674 ATTACTGTGCGCAACCGGAGATTTGCGCCGCGATTCGCGAA 1718
Db 1573 ATTCTATTTCAACGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1617

RESULT 5
ACF67367_48
Continuation (49 of 57) of ACF67367 from base 480001 (Photobadus luminescens nucleotide)
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
WP Fragment Name Begin End
WP ACF67367_01 1 11000
WP ACF67367_02 10001 21000
WP ACF67367_03 20001 31000
WP ACF67367_04 30001 41000
WP ACF67367_05 40001 51000
WP ACF67367_06 50001 61000
WP ACF67367_07 60001 71000
WP ACF67367_08 70001 81000
WP ACF67367_09 80001 91000
WP ACF67367_10 90001 101000
WP ACF67367_11 100001 111000
WP ACF67367_12 110001 121000
WP ACF67367_13 120001 131000
WP ACF67367_14 130001 141000
WP ACF67367_15 140001 151000
WP ACF67367_16 150001 161000
WP ACF67367_17 160001 171000
WP ACF67367_18 170001 181000
WP ACF67367_19 180001 191000
WP ACF67367_20 190001 201000
WP ACF67367_21 200001 211000
WP ACF67367_22 210001 221000
WP ACF67367_23 220001 231000
WP ACF67367_24 230001 241000
WP ACF67367_25 240001 251000
WP ACF67367_26 250001 261000
WP ACF67367_27 260001 271000
WP ACF67367_28 270001 281000
WP ACF67367_29 280001 291000
WP ACF67367_30 290001 301000
WP ACF67367_31 300001 311000
WP ACF67367_32 310001 321000
WP ACF67367_33 320001 331000
WP ACF67367_34 330001 341000
WP ACF67367_35 340001 351000
WP ACF67367_36 350001 361000
WP ACF67367_37 360001 371000
WP ACF67367_38 370001 381000
WP ACF67367_39 380001 391000
WP ACF67367_40 390001 401000
WP ACF67367_41 400001 411000
WP ACF67367_42 410001 421000
WP ACF67367_43 420001 431000
WP ACF67367_44 430001 441000
WP ACF67367_45 440001 451000
WP ACF67367_46 450001 461000

Db 66853 TTGTGTGGGTGTACGTTTGTGATGTCGTAAACCATATATCTGGAGAAATATTGCCCC 66912

Qy 1021 CGCGCAAAAGTAATTCACATTGATCATGACCGCGCGAGGTGGGAAAAAACAAGGCTCCC 1080

Db 66913 GAAGCAACGGATTACATATTGATTTGATCCAAATCAATTCGAAAAACGTTTCGGA 66972

Qy 1081 GATGTGCCCATTTGTGGGGATGTACGCCCATGTTTATGAAACAGCTTTTGACGCGGCCCCG 1140

Db 66973 GATGACTTATGTTGTGGATGCTAAACAGTGTCTTGAAATATGTTGTGTCGGAG 67033

Qy 1141 GATTTGGATTACCCCAACCATCCCATACACACCGGCGATGTTAAATGCGATTGATCAT 1200

Db 67033 ACAGTAGAAGACACACAAAGATCCAGATGCCCTAAAGATGTGGTTATCTATTACGCA 67092

Db		67093	TGGCGAGCCCTTAATGCTTTAAGTTTGATGCAGTGTGGTAAAGAATTAAACCAAGG	67155
Oy		1261	GNACTAACGGAATTTGGTCGCCAGC---CCCCGATGCCCTACTACAACCAGATGTGGGA	1317
Db		67153	GCTATCGAGGTTCTCTATCGTCTGACGAGGAGGAAGCTTATGTCACTTCAAGTGTGGG	67212
Oy		1318	CACACACAAAATGTGGGGGCCCAAGTTTTTGAACAATG---CCCCCGCATGAGATTCC	1374
Db		67213	CAGATCAAGATGTTTGTGATTCGATTCACCACTTTGATTAACCAAGGCGTTGAGTCACT	67272
Oy		1375	AGTCGTGCTTGGGTACGATGGGCTTTGTTTCTCCGCCCATGGAGCCCAAAGTG-G	1433
Db		67273	TCCGATGGTTTGGGGAGCGATGGGATTTGTTTACGGCGGCTTTGGCGCTTAACTGGTG	67332
Oy		1434	AGTGGGGGACGAGCGGCTCATTTTGCATCAGTGGAGATGCCAGCTTCCAAATGAACTTCAG	1493
Db		67333	A AACCGAGCGACGAGTAGTCTGTGTATTCGAGGGGAGTGTGATTAAGATGAATTTACG	67392
Oy		1494	GAACGTGGAAACCTTAGCCACAGTACGACATCCAGGTTAAAACTATTATCTCAATAACGCT	1553
Db		67393	GAGCTGTCCAACCGCATTCGATGCAATATGGGTGCCCAGTACTGGTTGAACTGAATAATCGT	67452
Oy		1554	TGGCAGGGGSAATGTGGGCTCAGTGGCAACAACCTTCTAGAGAAAGCTATCTGCTTCT	1613
Db		67453	TTCTTGGGAGATGATTAACACATGACAGATATATTTATTTAGCTGTGTCACCTCAATCT	67512
Oy		1614	AACATGTCCAGAGGCATGCCAGACATTAATCTCTCTGTGAAGCCATAGCATCAAGGCT	1673
Db		67513	TATATG---GAATCCTTCTGCTGATTTTGTTCAAATTTGGCTGAAGCTTACGGTCATGTGGC	67568
Oy		1674	ATTACTGTGGCGCAAGCGGGAGSATTGGCCCCCGCATGGCCGAA	1718
Db		67570	ATTCTTATTCAACGCTTATGATGAAATGGAAAATTAATTTGGCCGAA	67614
RESULT 6				
ACF65387_0				
WP Sequence split into 7 fragments				
WP	Fragment Name	Begin	End	LOCUS ACF65387 Accession Acf65387
WP	ACF65387_0	1	110000	
WP	ACF65387_1	100001	210000	
WP	ACF65387_2	200001	310000	
WP	ACF65387_3	300001	410000	
WP	ACF65387_4	400001	510000	
WP	ACF65387_5	500001	610000	
WP	ACF65387_6	600001	696798	
ID	ACF65387 standard; DNA; 696798 BP.			
XX				
AC	ACF65387;			
DT	20-NOV-2003 (first entry)			
DE	Photorehabdus luminescens nucleotide sequence #40.			
KM	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis; detection; food; gene expression; plant; animal; microorganism; tissue			

OY	421	GTGCTGATTTACGAGAGGGTGGCCGGGCGCATGATTTGGTGGCAGATGGCTTCCAGGAAT	480
Db	5826	GTGTGATTTGTCGGGTCAAGTCGCCAGTTGCTGTGATCGGTATGATGATTCCTCAAGATGT	5885
OY	481	GACATTTTGGCATCACTTACCGATGCTTAAGCACTCCTATGTGTGATGCTAGTGGCGG	540
Db	5886	GACATGTAGGAGATATCCCGCCGATGCTAAACACAGTTTTTTAGTGAATAAAGCAGAG	5945
OY	541	GATATGGCTCGCATTTGTTATCTAGAGGCTTTCACCTGCTGTAAGCACCCGGTGGTCCCGGCGG	600
Db	5946	GATATTCGGAAGGTACTGAAAAAGCTTCTTATTACTTCCTCAGTGGTCTCTCGGAGCC	6005
OY	601	GTTTGTGATCGATATTTCCCAAGAGTGTGGGCTTAGAAGATGTGAGTACATTTCCCTTGAC	660
Db	6006	GTCTGTATTGATTACCAAAAGACACGGTAAATCCGCGCATTAATAATACCTATATTAT	6065
OY	661	CCGGGTACGTTAATCTACCGGGGTATTCGCCGCCAGGTTAAAGGTATATCCCGCAAT	720
Db	6066	CTGTATCAGATAGATATGCTGCTTATTAATCAACGATTAAGGACATAAAGGCAAT	6125
OY	721	AATCGGCAATTTGCAATTTGTTGAGCAGCAGCCAGAAATCCCTTGCTCTACGTAGGGGAGG	780
Db	6126	AAGCGGCATTTAAGAAAACATTGTAATGCGAAAAACACATCATATATGTAGGGGGGGG	6185
OY	781	GCGATCGCCCGCATATGCCCATGCCAGGTGACAGAAATTTGCGGAAAGGTCCTCAGTGGCG	840
Db	6186	GCGATTACCGAGATTGTTCAAGCAAAATTCTGTATTTGGAGAAAAATTCATATCTCT	6245
OY	841	GTAACAACCAACCCGATGAGGAAATTTGGGGCTTTTGACCAAAAACATATCCCTTTGGTGGGT	900
Db	6246	GTGTGTAGCTCCCGATGGGCTTAGCGAGATTTCCCTGCACTCATTCGCAAGCCTGGCG	6305
OY	901	ATGTTGGGTATGATGCGCACCGGTATGCCATCTTTGGCCGTACAGCAATGTGATTTGTG	960
Db	6306	ATGTCTCGTATGACATGTGTACATTCGAAGCCAAACCAACCATATTTCAATTCAGATGAT	6365
OY	961	ATTGCAATGGGGGCGCCGTTTTCAGACGACCGGGTAACTGCAACATPAGACGATTTGCTAGC	1020
Db	6366	TTTCTGTCTGGTGAACGTTTGTGATGATGTAACCAATATCTGAGAAATATTGCGCG	6425
OY	1021	CGCCCAAGTAAATTCACATTTGACATGACCCGGCGGAGGTGGAAAAACAGGGCTGCC	1080
Db	6426	GAAACCAACGATTTACATATTGATATTGATTCACCAATCAATTCGAAACGGTTTCCGCA	6485
OY	1081	GATGTGCCAATTTGGGGGAGTACGCGCATGTTTAAACAAGCTTTTGACGCGGCGCCGG	1140
Db	6486	GATGTACCTATTTGTGTGATGATGTAACAGGTGCTTGAAACAAATGTTGTGTCGAG	6545
OY	1141	GAATTTGATTAACCCACCCATCCCATPACACCCAGGATGTGTAAATTCGATTTGATCAT	1200
Db	6546	ACACTTAGAAGCACACAAAGATCCAGATGCGCTAAACAAATGTGTGTGTTATCTATTACCAA	6605
OY	1201	TGGGCGACCGGATTAACCCCTCCAGGTGCCCACTATAGGATACTATTGGCCCCCAGAG	1260
Db	6606	TGGGCGGACCGGTAATATGTCTTAATGTTTGTATGCACTGTGTAAAGATTTAAACCAAGCG	6665
OY	1261	GTAAGTACCGAAATTTGTGCGCAGGC---CCCGATGCTTACTACCAACCGATGTGGGA	1317
Db	6666	GCTATCGAGGTTCTCTATCTGTCTGACCAAGGGAAMAACCTATGTCACTTCAGATGTTGGG	6725
OY	1318	CAACACCAAAATGTGGGGCGGCCAGATTTTGAACATGG---CCCGCCCGATGATTTCC	1374
Db	6726	CAGCATCAGATGTTTGTTCATTTGATTCATTAACCATTTGATTAACCAAGCGTTGATCAGT	6785
OY	1375	AGTGTGGCTTTGGGTACGATGGGCTTTGGTTTACTGTCCGCGCATGAGGACCAAAATGGG-G	1433
Db	6786	TCCGTGTGTTTGGGGAAGATGGGATTTGGTTTACCGCGCGCTTTGGGGGTTAACTGGTG	6845
OY	1434	AGTGGGGGACAGCGGTCATTTTGATCAGTGGAGATCCAGCTTCCAAATGAATCTTGAG	1493
Db	6846	AAACCGAAGGACGATGATCTGTGTTAACGGGGGATGTATATTCAGATGAATATTACG	6905

D	b	814	ACCGGCAATCCCGGAGGTGATTAGCAACGTGTATGSCACAGTGGCCGGGTTTCCCGACAGGCATCAC	873
O	y	889	CTTTCGGGTGGTATGTTGGTATGCAATGSCACCGCCTATGCAACTTTGCCGTACGCCGA	948
D	b	874	CAGAACTTGGGCAATGCCCGGCGATGACCGGCAACCGTTGGCGGCGTGGGCAACCGCTCGACGCCG	933
O	y	949	TGTGATTTGTTGATTTGACGTGAGGGGGCCGGTTTTCAGACGACCGGAGTAACTGCGAAACTAGAC	1008
D	b	934	AGCGATTTTACTGATTCGACTGCGACTGGGCACTGCTTTTGAACGACCGGGTGAACCTGGCAACTGCAC	993
O	y	1009	GAATTTGCTAGCCCGCGCCAAAGTAATTCAATTGACATCGACCCGGCGGAGGTGGAAAA	1068
D	b	994	TCTTTCGGGCCCGGACGCGCAAGGTTAATGATCCCGACATCGACCCGGCGGAGTGGCAAG	1053
O	y	1069	AACAGGAGTCCCGGATGTCGCCATGTTGGGGGATGTAGSCCAGTGTTTTGAACAGCTTTG	1128
D	b	1054	AACCGGACGCTGACGTGCGCATTTGTGGGCGACGTCAAAAGCTGTTATTTGATGAGCTCATC	1113
O	y	1129	CAGCGGCGCCCGGAGATTGGAATTACCCCAACCCATCCCATACACCCAGCAGCTGTTAAT	1188
D	b	1114	GCATGTGCTGCGCATTAAGAGGTCCCGGCAACATCGAAATGACGCACTGGTGTGCATAT	1173
O	y	1189	CGCATGTATCATTTGGCCGGAACCGATTACCCCTCCAGGTGCCCCACTATGAGGATATCTAT	1248
D	b	1174	TTGACGGGTGTTGGCAACGATATCCGTTGAGCTACAGTCAACAAAGCGAGCGCATCTGTG	1233
O	y	1249	GCCCCCGGAGGTATGACACAGAAATTGCT--CGCCAGCGCCCGCATGCTTACTACAC	1395
D	b	1234	AGCCCGGAGTACGTATTGAAAGCTTGTGTAATTTGTGGCCCGGAGCGGTATATGTC	1293
O	y	1306	ACCGATGTGGACCAACCAATATGTGGCGGCCACT--TTTGAACATATGGCCCCCGC	1362
D	b	1294	GCGGTTGATGATCAGCATCAGATGTGGCGCGGCGAGTTCATCTCTACGAAAGCCGGGT	1353
O	y	1363	CGATGGAATTTCCAGTGTGCTGGGTGATGAGTGGGCTTTGGTTTACCTCGCCGCAATGGGA	1422
D	b	1354	ACTTGGCTCACTCAGGCGGCTGTGGGACCAATGGGTTTGTCTATTCCGGCGGCATGGGC	1413
O	y	1423	GCCAA-AGTGGGAATGGGGGAGCAGACGCTCATTTGCAATCAGTGAGATGACAGCTTCAA	1481
D	b	1414	GCCAAAGATGTGCTCGCCCGGAGGACAGAAATCTGGGCGCATCGACCGGCGATGGCTGCCAG	1473
O	y	1482	ATGATCTTTACAGAACTGGGAACCTTACCAGTACATTCAGATTCAGAGTTAAATCTATTT	1541
D	b	1474	ATGACCAACAGGAGGATTGGCTACGTGCTATTCAGGGGCGCGCGCATCAAGTGGCGTTG	1533
O	y	1542	CTCAATAACGCTTGGCAGGGAGTGTGCGTCAGTGGCAACAACCTTTCTACAGAAACGT	1601
D	b	1534	ATCAACAACGCGCAACCTGGGCAATGTGTGGCAATGGCAAGCCTGTTTATTCAGGAGCGCT	1593
O	y	1602	TATTTGCTTTCAACATG-----TCCAGGCGCATGCCAGACATTAACTCTCTGT	1652
D	b	1594	TACTCGACAGCCGACCTGCGCCACCCATCAACCGCATTCGGAATTTGTGAAGTTGGCC	1633
O	y	1653	GAAAGCTATGGGATCAAGGATATTACT--GNGCGCAAGCGGGAAGATTTGGCCCCGCGG	1709
D	b	1654	GAGGCGTTGGGTGGCTGCGGATGGCTTGCAGATGTGAAGAGACGTTGTGATGTGATC	1713
O	y	1710	ATCGCCGAATAGTACCCCAACATGCTCTGTGTGATGATGTGTGTGCTCAAAAAGAT	1765
D	b	1714	AATCAGGCGACGGGCGCATCAACAACGCGCTGTGGTGAATCGACTTATGTTGTGGCGCGCAT	1773
O	y	1770	GAAACTGTACCTTACATGATATGGCCCCCGGCAATGATATAGGCCAAATGCTAGCT	1823
D	b	1774	GCGCGGTGTGGCGCATGCTGTGCGCTGGGCGCATCAACGACGAGATCCAAAGCT	1827

DT	19-JUN-2003	(first entry)
XX	Prokaryotic essential gene #19347.	
DE	Antisense; ds; prokaryotic essential gene; cell proliferation;	
KW	drug design; gene.	
XX		
XX	Mycobacterium avium.	
OS		
XX	WO200277183-A2.	
EN		
XX	03-OCT-2002.	
PD		
XX	21-MAR-2002; 2002W0-US009107.	
PF		
XX	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948893.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362659P.	
XX		
XX	(ELIT-) ELITRA PHARM INC.	
PA		
XX		
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;	
PL	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
XX		
DR	WPI; 2003-029926/02.	
DR	P-PSDB; ABU33820.	
PT		
XX	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX		
XX	Claim 14; SEQ ID NO 25560; 1766bp; English.	
XX		
PS	The invention relates to an isolated nucleic acid comprising any one of	
CC	the 6213 antisense sequences given in the specification where expression	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
CC	encoding a polypeptide whose expression is inhibited by the antisense	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC	polypeptide or its fragment whose expression is inhibited by the	
CC	antisense nucleic acid; (4) an antibody capable of specifically binding	
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
CC	proliferation or the activity of a gene in an operon required for	
CC	proliferation; (7) identifying a compound that influences the activity of	
CC	the gene product or that has an activity against a biological pathway	
CC	required for proliferation, or that inhibits cellular proliferation; (8)	
CC	identifying a gene required for cellular proliferation or the biological	
CC	pathway in which a proliferation-required gene or its gene product lies	
CC	or a gene on which the test compound that inhibits proliferation of an	
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a	
CC	compound's activity; (11) a culture comprising strains in which the gene	
CC	product is overexpressed or underexpressed; (12) determining the extent	
CC	to which each of the strains is present in a culture or collection of	
CC	strains; or (13) identifying the target of a compound that inhibits the	
CC	proliferation of an organism. The antisense nucleic acids are useful for	
CC	identifying proteins or screening for homologous nucleic acids required	
CC	for cellular proliferation to isolate candidate molecules for rational	
CC	drug discovery programs, or for screening homologous nucleic acids	
CC	required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> ,	
CC	<i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence is one of the target	
CC	prokaryotic essential genes. Note: The sequence data for this patent did	
CC	not form part of the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX		
XX		
XX	Sequence 1866 BP; 313 A; 673 C; 624 G; 256 T; 0 U; 0 Other;	
XX		
XX	Query Match 19.4%; Score 379.8; DB 7; Length 1866;	
XX	Best Local Similarity 54.3%; Pred. No. 1.8e-109;	
XX	Matches 938; Conservative 0; Mismatches 757; Indels 34; Gaps 7;	

QY 109 CTGAAAACCTGTTCTAGCCCAACGCGAAACCGGGCGTTATATCTGATGATAGCTTGAAA 168
Db 100 CCGAAACGCAATGGGCGCCGACAAAGTTATGTGTGCAAAATCGGTATCCGATCGCTGGAG 159
QY 169 CGCCAAATGGGGTCAACACATTTTGTCTATCCCGGCGGGGCAATTTTGCCATCTATATAT 228
Db 160 GAGCTCGGGGTGAGGTGATCTTCGCAATTCGCGGGCGGGTCTCGGCTCTATATAC 219
QY 229 GAATGTACCGCTTTGAAACGGCGGGGAAATTTGACATATTTTGTGCGCCATGAACA 288
Db 220 CGCTGTTGACT-----CAAAGACGTGGGCAAGTGTGTGTCCGCCACAGACAG 270
QY 289 GGAAGTTCCTCATGCGCGGATGGGTATGCAAGCCCAAGTAAAGTGGAGTTGTTTC 348
Db 271 GGTGCGCGATCAGCGCGCCACGCGCTACGCGCAAGCCACCGGCAAGGTGAGTGTGCAATG 330
QY 349 GGTATCATCTGACCAAGGGGCACTAATCTGTGACCGGATTTGCCAATGCCCATTTTGGAC 408
Db 331 GCCACTCCGATCCGGGCGCCACAACTGTGTACCGGCTGCGGACGCCCAAGTGAAC 390
QY 409 TCGGTGCCCATGTGTGATTACTGGAAGGTGGGCGGTGCCATGATGTAGCGATGCT 468
Db 391 TCGATCTCCGCTGTCGCGCTCACCGGCAAGGTGGAGCGAGCTGATCGGCAACCGACGC 450
QY 469 TTCCAGGAATTTGACATTTTGGGCATACCTTACCGATCGTTAAGCACTCCATATGTGTA 528
Db 451 TTCCAGAGAGCCGACATCTCCGCAATCCGATGCCATACCAACCAACTTCTGTGTC 510
QY 529 CGTATGCGGCGGCAATGAGCTGCAATTTTACTGAGGCTTTCATCTTGTAGCACCGAT 588
Db 511 CGCTCCGGACGAATCCCGAGGTGCTTCCGAGGGTTCCATATGCGCGCTCCGCGC 570
QY 589 CGTCCCGGCGCGGTTTGTATGATATTTCCCAAGATGTGGGCTTGAAGAAATGTGATGAC 648
Db 571 CGCCCGGCGGCTGCTCGTGCACATCCCAAGAGCGTGTGCAAGGGCGAGTGCACGTCCT 630
QY 649 ATTCCCTCGACCCCGGTGACGTATATCTAACCGGGTTATCGCCCAAGGTTAAAGGTAT 708
Db 631 AG-----CTGGCGCGCGGATCCACTGCGCGGATCAAGCCGACCAACCAACCGAC 684
QY 709 CCCCCAACAATTAATGCGGCATTTGCAATTTTGAAGCGCCAGAAATCCCTGTCTTAC 768
Db 685 AGCCCGCAGATCCGGAAGCCGCCAAGCTGATCGCGCGCGCAACCGATGTGTATC 744
QY 769 GTAGGGGGAGGGGCGATGCGGCCAATGCGCCATGCGCGAGGTGACGAAATTTGCGGAAAG 828
Db 745 GTAGGCGGCGGTGTATCTCGCGGAGGCGACGAGCTGTGCGGAATGCGCGAGTGTG 804
QY 829 TTCCAGTTGCGGTAAACAACACCTGTATGGGAATTGGGGCTTTTGAAGAAACCATCC 888
Db 805 ACCGCGATCCCGGTGTGACACGCTGATGGCGCGCGGCGGTTCCGGAACGCGACGCG 864
QY 889 CTTTGGGTGGTATGTGGGTATGATGCGCACCGCTATGCGCACTTTGCCGTACGCCAA 948
Db 865 CAGCACTTGGGATCCCGGATGACGCGACCGTGTGCGGGGTGGCGCGCTGTGAGCGC 924
QY 949 TGTGATTTGTTGATGCAATGGGGGCGCGTTTCAAGACCGGGTAACTGCAAACTAGAC 1008
Db 925 AGCGACCTGTGTATGCGCTGTGGTACGCGGTTTCAATGACCGGTGACCGGCAAGTGTAC 984
QY 1009 GAATTTGCTAGCCGCGCAAGTAATTAATTAATGACATGACACCGGCGAGGTGGAAAA 1068
Db 985 ACCTTGCGCCCGAAGCCAAAGTCAATCCACGCGCATGACCCGCGCAAGATCGGCAAG 1044
QY 1069 AACAGGGCTCCCGATGTGCCCATTTGTGGGGGATGTAAGCCATGTTTGAACAGCTTTTG 1128
Db 1045 AACCGGCAATGCGGATGCTGCGCATGTGCGGCAAGTCAAGGCGGTATGCGGAACTGTTC 1104
QY 1129 CAGCGGGCGCGGAATGTGATTAACCCCAATCCCATACCAACCGAGCATGTGTTAAT 1168
Db 1105 GAGATGCTGCGCAGAGGCGGCGCGCAATCTGACATGTGCGCATGTGTGGGCTTAC 1164
QY 1189 CGCATGTATCATTTGGCGAGCATTAACCCCTCCAGGTGCCCACTATGAGAGTATAT 1248

Db 1165 CTGACGACGTCCTCAGTCCACTATCCGCTAGCTACAGGGCGGAGAGCGGACGCGCTC 1224
QY 1249 GCCCCCAAGAGGTAGTACACGAATTTGTGCGCAGGCC--CCCATGCTCTATACAC 1305
Db 1225 GGGCCGATGACGTATCGAAGAGCTGGAGCAATGCGCGGCGGACGCGCTGTAGTGTG 1284
QY 1306 ACCGATGTGGGAACAACCAATGTGGGCGGCCAGT---TTTGAACAATGAGCCCGCG 1362
Db 1285 GCGGCGGTGTGATGATGATGATGTGGCGGCCCAATTCATCTCTACGAGAAACCGCGC 1344
QY 1363 CGATGATTTTCAAGTGTGCTTGGGTGATGATGATGATGATGATGATGATGATGATGAT 1422
Db 1345 ACCGTGCTCAATTCGCGCGCGCTGGGCAATGAGGATGAGGATGAGGATGAGGATGAGG 1404
QY 1423 GCCAAGTGGG-AGTGGGGGAGAGAGGATGATTTGATGATGATGATGATGATGATGATGAT 1481
Db 1405 GCCAAGTGGGCGGTGCGGAGCGCGGAGGTGAGGCGATGACGAGCGGATGATGATGATGAT 1464
QY 1482 ATGAATCTTGAAGAACTGGGAACCTTACCCAGTACAGATCCAGGTTAAACTATATAT 1541
Db 1465 ATGACCAACGAGAGTGTGCGCATCTGGCGCGGTGAGGAGATCCCGATCAAGGATGAGCG 1524
QY 1542 CTCAATPACGGTTGAGAGGGGATGTGCGGTGATGATGATGATGATGATGATGATGATGAT 1601
Db 1525 ATCAACAACGAGCACTGGGCAATGTGTGCGCAATGTGAGGAGCCTGTGTTACGAAAGCGG 1584
QY 1602 TATTC-----TGTCTTAAATGTCCAGGGGATGCGACAGATTAATCTCTCTGT 1652
Db 1585 TATTCGAAACGAGATGTGCGCACCACTGCGACCGGATCCGATCTTGTCAAGCTGCGC 1644
QY 1653 GAAGCTTATGATCAAGGAT---TACTGTGCGAAGCGGGAAGTATTTGGCCCGGCG 1709
Db 1645 GAGCGGTGTGGCTGTGCGATGTGCGATGTGCGATGTGCGATGTGCGATGTGCGATGTGCG 1704
QY 1710 ATGCGCGAATGTGTAAGCCCAATGATGCTGTGTGATGATGATGATGATGATGATGATGAT 1769
Db 1705 AAGCTGCGCGGCGATCAACGACCGGCGGTGTGATGATGATGATGATGATGATGATGATGAT 1764
QY 1770 GAAACTGTATCCCTATGATGTGCGCGGCGATGATGATGATGATGATGATGATGATGAT 1818
Db 1765 GCGCAGGTGTGCGGATGTGCGCGCGCGGACGACGACGACGACGATGATGATGATGATGAT 1813

RESULT 9
ACAS3435
ID ACAS3435 standard; DNA; 1722 BP.
XX
AC ACAS3435;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #35092.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Vibrio cholerae.
XX
EN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX MPI; 2003-029926/02.
DR P-PEDB; ABU49565.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 41305; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcc_sequences

XX Sequence 1722 BP; 471 A; 374 C; 451 G; 426 T; 0 U; 0 Other;

XX Query Match 19.4%; Score 379.6; DB 7; Length 1722;
XX Best Local Similarity 53.9%; Pred. No. 2e-109;
XX Matches 917; Conservative 0; Mismatches 764; Indels 19; Gaps 6;

QY 128 AACGGCAAAACCGGGCTTATCTCTGATGATACCTGAAAGCCATGGGGTCAACACA 187
DB 5 AAATGCTATCTGGCGCAGAGATGATCGCAATCTGATCAATGAAGGTGTGAGCAAA 64
QY 188 TTTTGGCTATCCCGGGGGGCAATTTGGCCATCTATGATGAAGTGAACGCTTGAAG 247
DB 65 TCTTGGATATCCCGGGGCTTCAAGTCTGACATCTACAGATGCCCTTCAAC-----GAAA 118
QY 248 CGCGGGGGGAAATTGAGCATATTTTGTGCGCCATGAAACAAGAGCTTCCCATCGGCGG 307
DB 119 AAACAGACCAATTAACACGTTTGGTTGTCAGAAACAAGCGCAACACATATGCTG 178
QY 308 ATGGGTATCCAGAGACCAAGTAAAGTGGAGTTTGTGGTACATCTGACACAGGCG 367
DB 179 ATGGCTATGCGGCGCCACTGAAAAACGGGGCTGAGTCTGTGCTCCAGGTCCGGGCG 238
QY 368 CGACTAATCTGGTGACCGGCATTTGCCAATGCGCAATTTGSACTCGGTGCCATGGTGTGA 427
DB 239 CCACCAATACGCTGATCTGGAGATTGCAACCGCTATATGACTGATTCAGATGATTGTA 298
QY 428 TTACTGAGAGGTGGGCGGTGCATGATTGGTAGCATCTTTCCAGAAATTACATTT 487
DB 299 TCTCGGTACGTAGCCCACTAATCTTATTTGTAAGATGACATTTCAAGAGTGCATATCG 358

QY 488 TTGGCATCACTTACCGCATCGTTAAGCACTCCTATGTGTGACGTAGTCCGCGCATATGCG 547
DB 359 TCGGGGTGTCTGTCCCATCTGTTAAACAAGCTTTTGTGTAAAAAACCACAAATATTC 418
QY 548 CTCGATGTGTACTAGAGGCTTCCATCTGTGCTAGACACGGTCTGTCCGGCGGTTTGA 607
DB 419 CTGAAACCATCAAAAAAGCGTTTATATCGCTCAACGGGGCGTCTGTGTTGATGAGA 478
QY 608 TCGATATTTCCAAAGATGTGGGCTTGAAGAAATGATGACATTTCCCTCAACCCCGGTG 667
DB 479 TTGACGTGCAAAAATGATGATGAGCCCGCTTAAATACCTGCTTATGATGCCAATA 538
QY 668 AGCTTATATCTACCGGGTATATGCGCCAGGTTAAAGATATCCCGCAAAATTAATGCGG 727
DB 539 CATTAAATATGCTCTTAAACCAACAGACGCGCTGCGCAAGAGACATCAAAAAG 598
QY 728 CATTCGATTTTGGAGACAGGCGCAAAATCCCTTGTCTAGCTAGAGGGAGGCGGATCG 787
DB 599 GGTATCGGCACTGTCTCATGTGCAAAAAAGCCAGTACTAGCTGAGTGGCGGGCGATTA 658
QY 788 CGCCCAATGCCATGCCACAGGTGACGAATTTGGCGAAAGATTCAGTTGCCGATCAAA 847
DB 659 TTGCTAATGACATGACGAATTTGCCCAACTTGGACGGCTTAAATCTGCCCGGTTA 718
QY 848 CCACCCGTATGGAATTTGGGCTTTTGAAGAAACATCCCTTGGTGGATATGTTGG 907
DB 719 GTACTTTATGATGGATTTGGGCGTTTCTCGGACATCAATAAAACCTCACTGGGTATGCTG 778
QY 908 GTATGATATGCGCACCGCTATGCGCACTTTGCCGTAGCGAATGTGATTTGTGATGAC 967
DB 779 GCATCATGTGTGTATGATGAAGCAATATGGGATGACAAATGCAATCTGATTTTGGGG 838
QY 968 TGGGGGCGCGTTTGCAGACAGCGGATGACGAACTAGACGAATTTCTAGCCGCGCA 1027
DB 839 TAGGTGTGCGCTTTATATACGCTATGACATTAACACTTGAAGATCTGCCCAAGCGCA 898
QY 1028 AAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1087
DB 899 AAATCATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 958
QY 1088 CCATGTGGGGGATATGAGCGCATTTTGAAGCAAGCTTTTGCAGCGGGCGCGGAAATGG 1147
DB 959 CGATTTGTGTCTGCGCAGCAAGATTTGATGATGATGATGATGATGATGATGATGATG 1018
QY 1148 ATTACCCCAACCATCCCATCCATCCACCAAGCATGTGTAATTCATTCATTCATTCATTC 1207
DB 1019 CTGAGGCAATATGATGCGGAGCATTTGATGATGATGATGATGATGATGATGATGATG 1078
QY 1208 CCGATTAACCCCTCCAGGTGCGCCCATATGAGATACATTTGCCCGCAGGA---GATAG 1264
DB 1079 ATGCTAGTGTGTGCTTATGAACTCAGCAAGAAAGATTAAGCTTACGCAAGTATG 1138
QY 1265 TACAGCAATTTGTGCGCAGGCGCCGATGCTTACACCAACCGATGTGGACAAACAC 1324
DB 1139 AATCGTATATTAATTAACAGGAAGGTAAAGCAATTTTGGCGTGGATGTGGGGCAGCAC 1198
QY 1325 AAATGTGGCGGC---CCAGTTTGAACATGCGCCCGCGCATGATTTCCAGTGTG 1381
DB 1199 AAATGTGTGCGGCACTTCACTATTCGTTTGAACCGCGCTCAGAGATTAACCTGTGCG 1258
QY 1382 GCTTGGGTACATGAGGCTTTGTTTACCTGCGCCATGAGGAGCCAAATGTGGAAGTGGGG 1441
DB 1259 GGTATGACACATGAGTGTGTTTGTATCTGCGGAAATGGCGTGAATTCGCAATGCCAG 1318
QY 1442 ACG-AGCGCATTTTGCATCAATGAGATGCAAGCTTCCAAATGATCTTACAGAACTGG 1500
DB 1319 AAGAAGAGGTACTGTGTCGACCGGAGATGAGATTCAGATGATTAATTCAGAGGCTAT 1378
QY 1501 GAACCTATGCGCAGTACGACATCCAGGTAAACATTAATCTCATATACGATGTGGCAGG 1560
DB 1379 CGACGCACTGCAATGACGATTCATTCAGTAAATATTAATTAACCTCAACAAACGTTTCTTG 1438

QY 1561 GGATGTCGTCAGTGGGACAACTTCTACGAGACGTATTGCTTCTAACATGT 1620
 DB 1439 GAATGTGAAACAGTGGCAGATATTATTATCAAGGTCGTACCTTAACATACATGA 1498
 QY 1621 CCCAGGCAATGCCACATTAATCTCTCTGTGAAGCTATAGCATCAAGGATTAATCTG 1680
 DB 1499 GTTCTG---TGCCGATTTTGGCGGATCGAGAAAGCTTATGTCTATGTTGGATCCGTA 1555
 QY 1681 TGCCGACGCGGAAATTTGGCCCGGATCGCCGAATGCTAGCCACATGATGTCCTG 1740
 DB 1556 TTTCACACCTGATGAGCTGCAAGCTGGGCTGGAAGAACCTTAGCCATGAAAGATCCTC 1615
 QY 1741 TGGTGT---ATGATGTGCTGTCAAAAAAGATGAAACTGTATACCTATGATGTCCTCCG 1797
 DB 1616 TGGTGTGTCGACATTAATGTATGATGAAACTGAGCAGCTTACCCATGACATCAAG 1675
 QY 1798 GCATGATTAAGCCCAATG 1817
 DB 1676 GCGAAGGATGACAAATG 1695

RESULT 10
 ID ACA23942
 AC A23942 standard; DNA; 1875 BP.

AC ACA23942;
 DT 19-JUN-2003 (first entry)
 DE Prokaryotic essential gene #5599.
 XX
 KM Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX drug design; gene.
 XX
 OS Borrelia cepacia.
 XX
 PN MO20027183-AZ.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002MO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX P-PSDB; AB020072.
 DR WPI; 2003-029926/02.

XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 11812; 1766pp; English.

XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (?) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 1875 BP; 390 A; 591 C; 583 G; 311 T; 0 U; 0 Other;

Query Match 19.3%; Score 378.8; DB 7; Length 1875;
 Best Local Similarity 54.7%; Pred. No. 3.7e-109;
 Matches 938; Conservative 0; Mismatches 747; Indels 31; Gaps 8;

QY 136 ACCGCGCTTATATCCGTGATGATGACCTGAAAGCCATGCGGCTCAACACATTTTGGC 195
 DB 178 ATCGGCGGACACGCTGCTCATGAGACCTGCGCAGAAAGATGCAATTCATCTGGGG 237
 QY 196 TATCCCGCGCGGCAATTTTCCCATCTATATGAACTGTACCGCTTTGAAGCGCGGG 255
 DB 238 TACCCCGCGGCTCGGTCTCTACATCTACATGACAGACTGTAC-----AGCAGGAC 288
 QY 256 GAATTTGAGATATTTTGGGCGCCATGAAACAAGAGCTTCCCATGCGCGATGAGGTAT 315
 DB 289 AAGATTCAGACCTGCTGTGTGCCCATGAAAGCGCGCTGTACCCACCATGAGTAT 348
 QY 316 GCCAAGACCAAGATTAAGTGGAGTTTGTTCGTACATCTGACACGAGCGGACTAAC 375
 DB 349 GCCGCTCCACCGGCAACGTCGGCTCTGTGACGTCGCGGCCCGCGTGCACCAAT 408
 QY 376 TTGTGACCGGCAATGCGCATGCGCATTTGGAATCGGTGCCATGATGATTAATGCA 435
 DB 409 GCGGTACCGGCAATCGAAGCGCTTACATGATGATTCATCCATGCTGATCAGCGG 468
 QY 436 GAGGTGGCGCGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
 DB 469 CAGGTGCGGACCGGCGCGCATGCGTACGATGCGTTCCAGAAATGCGACACCGTCCG 528
 QY 496 ACCTTACCATGCTTAAAGCACTCTATGTGTACGTAAGTGGCGGATATGCTGCATT 555
 DB 529 AGCGTCCGTCGTGAAGCAACATTCCTCTGTAAAGAGCGTGGCACTCGGGAAC 588
 QY 556 GTTACTGAGGCTTTCATCTCTGTACACCGGTCGTCGCGCGCGGTTTGAATGATAT 615
 DB 589 GTCAAGAGCGCTTATATATGCGCCGACCGCGTCCGCGCGCGTGTGATGACATC 648
 QY 616 CCCAAGATGTGGGCTTAAAGAAATGTGATGATGATGATGATGATGATGATGATGAT 675
 DB 649 CCCAAGACATCTCGAAGACCGCGTCCAGTACGACCGCATCA-----GAGCGGTGC 702
 QY 676 CTACCGGATATGCGCCGACGTTAAAGTAAATCCCGCAATTAATGATGCGGATGCA 735
 DB 703 CTGCTGTGACCAACCGTACGAAAGGCGATTCGGGTTCAGATCCGACAGGCGGTG 762
 QY 736 TTGTTGAGACGCGCAAAATCCCTGCTTACGTAGTGGGAGAGGCGATCGCGCAAT 795
 DB 763 CTGCTGTGACGCGCAAGGCTCGTACATCTACACGCGGCGGCGATCTCTCGCGCAG 822
 QY 796 GCCCATGCCAAGGTGACGATTTGCGGAAAGGTTCCAGTTGCGGTAACACCAACCCG 855

Db 823 GCGTCGCTGAATGAAACCAAGTTGCGGACCTGCTCGGCTAACCCGGTCAAGAACACGCTG 882
 QY 856 ATGGGAATTTGGGCTTTTGAAGAAACCATCCCTTCGTGGGTATGTTGGTATGCAT 915
 Db 883 ATGGGCTCTGGGGGTATGCGCGTCCGGAACAAGAAATTCCTGGCATGCTCGGCATGCAC 942
 QY 916 GGGCACCCTGATGCAACTTTGGCGTCAAGCAATGTATTTGTGATTTGCAATGGGGGCC 975
 Db 943 GGCACCTACGAAGCAACATGCGATGCAAGCATGCGACGTCGATGCGATGCGGCGCC 1002
 QY 976 CGTTTCAGCAGACCGGGTAACTGAGCAAACTAGACGAATTTGCTAGCCGCGC---CAAGTA 1032
 Db 1003 CGCTTCAGACGACCGGTGATGGGGACCCGGCGACCTTCGGGTGCGCCCGCGCAATC 1062
 QY 1033 ATTCAATTTGATGATGACCGCGCGGAGGTGGAAAAACAGGGCTCCCGATGTCCTATT 1092
 Db 1063 ATTCAATTCGACATGACCGCGCTGCTCGATCTCGAAGCGCGTGAAGTTCGATCCCATC 1122
 QY 1093 GTGGGGGATATGACGATGTTTATGAACAGCTTTTGGACGGGGCGCGGAAATTTGATTA 1152
 Db 1123 GTCCGCGACGTGAAGGAAGTCTGAAGAGCTGATCGACAGCTTGACAGACGGCCGACAT 1182
 QY 1153 CCCACCCATCCCATACCAACCCAGGCATGTTAAATGCGATTGATTCATTTGGCGGACGAT 1212
 Db 1183 GGGCCCGACACCGAAGCGCTCGCCAAATGTGGAAGACATCGAAGGCTGGCGCGGAG 1242
 QY 1213 TACCCCTCCAGGTGCCCCCACTATGAGATCTATTCGCCCGCCGAGGATGATACAGAA 1272
 Db 1243 GACTGCTGAAGTACGATCGGAAGCGAATCATCTAAGCCGACGTACGTTGAGAAAG 1302
 QY 1273 ATTGTCGCCAGGCCCCCG--ATGCCCTACTACCAACGATGTGGGACAAACCAATG 1329
 Db 1303 GCGTGGAGCTGACGGAAGCGCAATGCGTTCGTGCTCGACGTCGCGCACACCAATG 1362
 QY 1330 TGGCGGGCCAGTTT---TGAACATGCCCCCGCGATGAGATTTCCATGCTGGGTTG 1386
 Db 1363 TGGGCGCGCATCTTACCTGTTTCAACAAGCGGCTGCTGATCAACTCCGGCGGCTC 1422
 QY 1387 GGTACGATGGGCTTTGTTTACTGCGCGCCATGAGAGCAAAAGTGGAGTGGGGAGCA- 1445
 Db 1423 GGCACGATGGGCTTCGCGCTGCGCGCGCGCATGGGCGTCAAGATGCGCCATCGGAGAC 1482
 QY 1446 GGGGTCAATTGATCATGATGAGATGCCAGCTTCCAAATGATCTTCCAGAACTGGGAAC 1505
 Db 1483 GACGCTGTGTATCATCAGGCGGAAGGCTCGATCCAGATGTGCATCCAGAACTGTCCAGC 1542
 QY 1506 CTAGCCGATGACGATCCAGATTAATACTATTTCTCAATACGTTGGCGAGGGATG 1565
 Db 1543 TGCTGCGATGACAGACGCCCTGTAAGATCTTTGCTGAACAAACGCTACCTCGGATG 1602
 QY 1566 GTGCGTCAAGTGGCAAACTTTCTACGAAGACGTTATCTGTTCTTAACATGTCCAG 1625
 Db 1603 GTCCGCGAGTGGCAGAGATGAAATACAGCAAGCGCTATTCGCAATGCTGTCATG---GAT 1659
 QY 1626 GGCATGCGACATTAATCTCTCTGTAAGCCATGAGCATCAAGGTTATTACTGTGCGC 1685
 Db 1660 GCGCTGCGGATTTCTGTGAAGCTGCGGAGCGTACGCGCATGTGCGCAATCGGATGAA 1719
 QY 1686 AAGCGGGAAGATTGGCCCCCGCGATGCGCGAAATGTAGCCCAATG---GTCTGTG 1742
 Db 1720 AAGACTTCGATGTGAGACCGGCGCTGAAGGAACGCTGGCGCTGAAGACCGGACGCTG 1779
 QY 1743 GTGATGATGTGGGTGTCAAAAAAGATGAACCTGTAACCTTAATGATGCCCCGGGATG 1802
 Db 1780 TTTCTGACTTCAGACCGATCGAACCAAAACGCTGCGGATGTACAGGCGCGGAG 1839
 QY 1803 AGTAATGCCAAATGCTAGTTTACCGGAAGTGGCG 1838
 Db 1840 GGCATACCGAGATGTCTCTCGATCGGAAGATCTG 1875

RESULT 11
 AAS54310

ID AAS54310 standard; DNA; 1725 BP.
 AC AAS54310;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Pseudomonas aeruginosa DNA for cellular proliferation protein #441.
 XX
 KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 KM antibacterial; drug design.
 OS Pseudomonas aeruginosa.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 XX
 PT P-PSDB; AAU36451.
 PT
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Claim 27; SEQ ID NO 7947; 511p; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 CC
 SQ Sequence 1725 BP; 351 A; 546 C; 517 G; 311 T; 0 U; 0 Other;
 XX
 Query Match 19.1%; Score 373.8; DB 4; Length 1725;
 Best Local Similarity 54.6%; Pred. No. 1.4e-107;
 Matches 866; Conservative 0; Mismatches 712; Indels 25; Gaps 6;
 QY 135 AACCGGGGCTTATATCTGATGATGAGCTGAACCGCCATGGGGTCAAAACATTTTGG 194
 Db 12 ATCTGGCGGTGAAGATGTGTCGCTCGTTCGCGACGAAAGCGCTTAAGTACATCAAG 71
 QY 195 CTAATCCGCGCGGGAATTTTGGCCATGATGATGAAGTGAACCGCTTTGAAGCGGGGG 254
 Db 72 GTACCGGGGTGTGCTCTCTGCAATATCAAGACGCTCTTC-----AAGAGCA 122

QY 255 GGAATTGAGCATTTTGGTGGCGATGAAACAAGAGCTTCCATGGGGCGAGTGGTA 314
 Db 123 GCACTGACCCACATCTGTTGGTGGCGACAGACGGCGGTACCTCATATGGCCGAGGCTA 182
 QY 315 TGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGCTACATCTGAGACGAGGCGACTAA 374
 Db 183 GCGCCGCGCCACCGGCAAGCCCGGTGTGTGCTGTGATCTTCGCGCCGGGGGGAGCA 242
 QY 375 CTGTGTGACCGGCACTTGGCAATGCCCATTGGAATCTGGTGGCCATGGTGGTATCTGG 434
 Db 243 GCGCATCACCGGTATCGCACCGCATACATGATGATCCCGATGGTATCTGTTCGG 302
 QY 445 AGAGTGGGCGCGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
 Db 303 CCAAGTGGCGAGCAACATGTGTGATCCGATCCGATCCGATCCGATCCGATCCGAT 362
 QY 495 CACCTTACCGATGCTTAAAGCACTCTATGTGTAGTGTGTGTGTGTGTGTGTGTGT 554
 Db 363 CTCCCGCGGATCGTGAAGCAAGCTTATCATCAAGATCCCTGGAAATTCGAGAGT 422
 QY 555 TGTATCTGAGGCTTTCATCTTGTAGCAACCGGTGTGTGTGTGTGTGTGTGTGTGT 614
 Db 423 GATCAAGAGGCTTCTTACCTGTGGCGAGTCCGTTGTGTGTGTGTGTGTGTGTGT 482
 QY 615 TCCCAAGAGTGTGGCTTAAAGAAATGTAGTACATTCCTCCGATCGACCCCGGTGA 674
 Db 483 TCCGAGAGCATGGGGGAGCCGACGCAAGAAATTCGATATTCCTATCCGAGAGGTCA 542
 QY 675 TCTACCGGCTTATCCGCGCCACGCTTAAAGTATCCCGACAAATTAATGGGCACTT 734
 Db 543 GTTGGCTGTGACACCGCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
 QY 735 ATTGTGAGCAGCAGGCAAAATCCCTGTCTACGTAGGGGAGGGGCGATGCGCGCA 794
 Db 603 GATGCTCTGGCGCGCAAGCGCCGCTGTCTATTCGCGCGCGCGCGATGTATGGGCA 662
 QY 795 TGCCCATGCGCAGTGTGCAAGAAATTTGCGAAAGTTTCCAGTTGCGGTTAACACCA 854
 Db 663 TGCCCGCGCGCGCTGACCGAGCTGCGAGATGCTCAACCTGCGCGGTGACCAACAC 722
 QY 855 GATGGGAAATGGGCTTTTGAAGAAACCATCCCTTTCGAGGATGATGTTGGGATGCA 914
 Db 723 GATGGGCTTCCGCGGATATCCGCGCGACACCGCATTCCTCGGCAATCTCGGATGCA 782
 QY 915 TGCGCACCGCTATGCTCAACTTGTGCGGTGACGCAATGTATTTGTATGTGAGTGGG 974
 Db 783 TGCGCAGCTTCAACCGCAACTGCGATGATCAGACGACGATGATCTCGCGTGGGCG 842
 QY 975 CGGTTTCAAGCAGCGCGGTATCTGCAAACTAGCAAAATTTGTAGCCCGGCAAGTAT 1034
 Db 843 GCGTTTCAAGCAGCGCGGTATCTGCAAACTAGCAAAATTTGTAGCCCGGCAAGTAT 902
 QY 1035 TCACATTGACATGACCCGCGGAGAGTGGGAAAAACAGGCGTCCCGATGTGCCATTGT 1094
 Db 903 CCACATGACATGACATCGCGGCTGATTTCCAAAGCAATCAAGGCGCATCCGATGCT 962
 QY 1095 GGGGGATGTACCCCATGTTTGAACAAGTTTGAAGCGGCGCGCGGAATTTGATTA 1154
 Db 963 CGGCGCGGTGACAGCGTTTCAACCGAAATGTGCGGATGTCAGAGGAAATTCGGGAG 1022
 QY 1155 CACCATCCCATACACCCAGCATGTTAAATGCAATGATCTATTGGC---GGACCGA 1211
 Db 1023 CCGGAACAGAGATGCCAGGCTGCTGTGTGAAGAGATCAGACGATGGCGGTGTAACG 1082
 QY 1212 TTACCCCTTCCAGTGTGCCCTACTATGAGATATCTATGCCCCCGAGAGTATACAGA 1271
 Db 1083 TGGGTGTGTTCCCGTACAGAAAGGCGACGCGCATCATCAAGCGCGAGACCGTATCAT 1142
 QY 1272 AATTGGTGGCGAGGCGC-----CGATGCTTACACCAACCGATGGGAGCAACACCA 1335
 Db 1143 GACGCTTACAGGTTTACCATGAGGATGCTTCACTCACTTCGATGTGCGGACAGACA 1202
 QY 1326 AATGTGGCGCGCCAGTTTGA---CAATGGCCCCCGCGATGATTTCCAGTGTGG 1382

Db 1203 GATGTTGGCTCCGAGTACTACAAAGTTCAACAGCCCATGCTGTGATCAATCCGCGG 1262
 QY 1383 CTGGGTACAGATGGGCTTTGGTTTACCTGTGCGCATGGAGCCAAAGTGGAGTGGGGA 1442
 Db 1263 TCTGGGACCATGAGGCTTGGCTTCCGCGCGAGATGGGATCAAGCTCAACTTCCGGA 1322
 QY 1443 CGA-GCGGTATTTGATCATGTGAGATGCGACCTTCCAAATGAATCTTCAAGAACTGG 1501
 Db 1323 CGAGATGTGCTGTGTGAGCCGCGCAAGGCAATATCCAGTAAATTCAGAACTGTG 1382
 QY 1502 AACCTTACCCAGTACAGATCCAGTTAAACTATTTATCTCAATPACGTTGGAGGG 1561
 Db 1383 TACCTGTGCTGACATGACACTGTGCGGTGAAGATGCTCACTGAACAAATGTGCTGG 1442
 QY 1562 GATGTGTGCTGATGCGACAAACTTTTACGAAGAAGTTATTTGCTTCTTACATGTC 1621
 Db 1443 CATGTGTGCTGCTGTGAGCCGCGCAAGGCAATGCAATCAACACCGTTATTCGATTC 1500
 QY 1622 CCAAGGATGCGAGCATTAATCTCTGTGTGAAGCTATGGCATCAAGGATTAATCTG 1681
 Db 1501 -GATCTCTCCGAGCTTGTCTCAAGCTTGTCCGAAAGCTTACGAGCATGTGGCAT 1559
 QY 1682 GCGCAAGCGGAAAGATTTGGCCCGCGGATGCGCAATGTCTAGCCACAAATGTCCT 1741
 Db 1560 CACGACCTGAAGATCTGAAGCCGAAGATGAGAGAGGCTTGGCATGAAGATGCGCT 1619
 QY 1742 GGT 1744
 Db 1620 GGT 1622

RESULT 12
 ACA42654
 ID ACA42654 standard; DNA; 1725 BP.
 XX ACA42654;
 AC
 XX
 XX 19-JUN-2003 (first entry)
 DT
 XX
 XX DE Prokaryotic essential gene #24311.
 XX
 XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX OS Pseudomonas aeruginosa.
 XX
 PN W0200277183-A2.
 XX
 XX 03-OCT-2002.
 PD
 XX
 XX 21-MAR-2002; 2002W0-US009107.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;
 PI Wall D, Twickie JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU38784.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 30524; 1766pp; English.
 PS

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the largest
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 1725 bp, 351 A; 546 C; 517 G; 311 T; 0 U; 0 Other;

Query Match 19.1%; Score 373.8; DB 7; Length 1725;
Best Local Similarity 54.6%; Pred. No. 1.4e-107;

Matches 886; Conservative 0; Mismatches 712; Indels 25; Gaps 6;

QY 135 AACCGGGCTTATTCCTGATGATAGCGCTGAAGAGCGATGGGTCAAACATTTTGG 194
DB 12 ATTCGGCGCTGAATGTCGTCGCTGTCGCGAGAGAGCGCTTAAGACATCTACGG 71
QY 195 CTATCCGCGGGGGAATTTTGGCCATCTATGATGAATGATGACGCTTTAAGCGGGGG 254
DB 72 GTACCCGGGTGTGCTCTCTCTGATATTAAGACGCGCTCTTC-----AAGAGCA 122
QY 255 GGAATTGAGCATATTTGTGTGCGCCATGAACAGAGCTTCCGATCGGCGGATGGGTA 314
DB 123 CGACGTGACCAATCTCGGTGCGCCAGAACAGCGGCTAACCATATGCGCGACGCTA 182
QY 315 TGCCAGAGCCACAGGTAAAGTGGAGTTTGTGGTATCATCTGGACAGAGGGGATTA 374
DB 183 CGCCGCGCCACCGGCAAGCCCGGTGTGTGTGTGTATCTCCGCGCGGCGGACCAA 242
QY 375 CTGTGTGACCGGCAATTCGCAATGCGCATTTGAGCTCGGTGCCATGGTGTGATTA 434
DB 243 CGCATATCACGGTATTCGCCATCATATGATCATCCCATGGTGTGATTCGTGGG 302
QY 435 AGAGGTGGCCGCTGCGCATGATTTGGATGATGCTTTTCAAGAAATTGACATTTTGGCAT 494
DB 303 CGAGTGGCCGACCAACATGTCGATCGATGCGTTCCAGAGAAACCAACATGTCGGGAT 362
QY 495 CACCTTACCGATCGTTAAGCACTCCTATGTGTGATGATGAGTGGCGGATATGGCTGCAT 554
DB 363 CTCCCGCGCATCTGTGAAGCAACAGCTTATCATTAAGCATTCCTCGAATTTCCCGAGT 422
QY 555 TGTATCTAGAGCTTTTCATCTGTCTAGACCGGTGTCGCGGCGGGGCTTTTGTATGATAT 614
DB 423 GATCAAGAGAGCTTTCTACCTGTGGCGAGTCCGCTGTCTCGGGGCGGTGTCTGTGATAT 482
QY 615 TCCCAAGGATGTGGGCTTAGAAGATGTAGATCATTCCTCGACCCCGGTGACGTTAA 674
DB 483 TCCCAAGGATGTGGGCGACCGGACGACGAGAAATTCGAATTTCTATCGAAGAGTCAA 542

QY 675 TCTACCGGGTATTCGCCACGAGTAAAGTATCCCGCAAAATTAATGGGCAATTGCA 734
DB 543 GTTGGCTTGTACACCCCGGCGGTTCGCGTACTCGGACAGATCCGAAAGCGCCGA 602
QY 735 ATTGTGAGCAGCGCCAGAAATTCCTTCTACCTAGAGGGAGGGCGATGCCGCCAA 794
DB 603 GATGCTCTCGGCGCGCAAGCGCGCGGTGTATTCGCGCGCGCGGTATCATGAGGCA 662
QY 795 TGGCATGCGCCAGGTGACAGAAATTTGGGAAAAGTTCCAGTTTCCGGTAAACACCTT 854
DB 663 TCCCGCGGCGCGTACCGAGAGCTGGCGAGATGCTCACTGCGCGTACCAACCTT 722
QY 855 GATGGAATTTGGGGCTTTTGAAGAAACATTCCTTCCGTGGGTATGTTGGATGCA 914
DB 723 GATGGGCTCGGCGGGGTATTCGCGGAGACCGGCAATTCCTCGGCATGCTCGGATCA 782
QY 915 TGGCAACCGCTATGCCAACTTTCGCTAGCGATGTGATTTGTGATGACATGGGGCG 974
DB 783 TGGCAGCTTACCGGCAACCTGCGCATCATCACGACGATGATTCCTCGGCTCGGCG 842
QY 975 CGTTTTCAGCAGCGGGTATCTAGCAAACTGCAATTTGCTAGCCGCGCAAGTAT 1034
DB 843 GGTTCGACAGCGGGTCTATACGTGCGGAGATTTGCGCGAAGCCGCAAGATCAT 902
QY 1035 TCACATTCATTCGACCGCGGAGGTGGAAAAACAAGGCTCCGATGTGCCATTTGT 1094
DB 903 CCACATTCAGATCATTCGCGGCTGATTTCCAGACCATCAAGCGCGCATCCGATCGT 962
QY 1095 GGGGATGACCCCATGTTTAAACACCTTTTGCAGGGGCGCGGAAATGATTAACC 1154
DB 963 CGGCGCGGAGACCGGCTTCTACCGAATGCTGCGATTCGTACGAGAAATCGCGAAC 1022
QY 1155 CACCATCCCATTCACCCACCGAGTGTAAATGATGATATTCATTCGCGCGGCG 1211
DB 1023 CCGAAGACAGATTCGCGGCTGCTGTGGAAGATTCAGACAGTGGCGGTGTAACG 1082
QY 1212 TTAACCCCTCCAGGTGCCCATATGACATATTTTCCCGCCAGAGGTATGACAGA 1271
DB 1083 TGGTGTTCCTCGTACGACAGGCGAGGAGCATCATCAAGCGCGACGCTCATGGA 1142
QY 1272 AATTGATGCGAGGCGCC-----CGATGCTTACTACACCAAGTGGGACAAACCA 1325
DB 1143 GACGCTACAGAGTATTCATGCGATGCGATGCTTATCACTTCGATGCGCGACAGCA 1202
QY 1326 AATGTGGGGGCGCCAGTTTGA--CAATGGCCCGCGGATGATTTCAAGTCTGG 1382
DB 1203 GATGTTCCTCGCAGTACTCAAGTTCAACAGCCCAATGCTGATCAACTCGCGCG 1262
QY 1383 CTGGGTACGATGAGGCTTGTGTTTACTGCGGCCCATGAGGCAAGTGGAGTGGGGA 1442
DB 1263 TCTCGGACATGAGCTTCGCTTCCGCGCGCATGAGGATCAAGCTCAACTTCGCGA 1322
QY 1443 CGA-GCGTCATTTGATCACTGATGAGATGCAAGCTTCCAAATGAATCTTCAGAACTGG 1501
DB 1323 CGACATGTCGCTCGTACGCGGCGAGGAGATTCAGATGAATATCCAGGAATGCTC 1382
QY 1502 AACCTTACCCAGTACGATTCGAGTAAATCTTAACTTATTCGATTAACGTTGGAGG 1561
DB 1383 TACCTGCTGCTGACGACCTGCGGTAAAGTGTCACTGAACATGAGTGGCTTGGG 1442
QY 1562 GATGTGCGTCAGTGCAACAACTTCTTACGAAAGAGTATTTCTGTTTAAATGTC 1621
DB 1443 CATGTTCGCGAGTGGGAGGACATGACATTAACAAGCGTTATTCGATTCCTCATG- 1500
QY 1622 CAGGGCATGCCAGCATTAATCTCTCTGTGAAGCTTATGATCAAGGATATTAAGT 1681
DB 1501 -GAATTCCTCGCGAGCTTGTCAAGCTTTCGGAAGCTTACGAGGATGTGGCATGCGAT 1559
QY 1682 GCGCAAGCGGAAAGTTTGGCCCGCGCGAGTGGCAAAATGCTAGCCCAATGATCTGT 1741
DB 1560 CACGACCTGGAAGATCTGAAGCGAAGATGAGAGAGGCGTTGCGCATGAAGATCGCT 1619

Db 1369 GGCACGCGGACGACGCTTCCGACCCCGGTGAGAGATTCCTGAAACAACGAGCGGCTG 1428
 Qy 1560 GGGATGTCGTCGTCAGTGGCAACAATCTTTCAGAGAGAGATTTCTGCTTTAAACATG 1619
 Db 1429 GGCATGTCGCGGACGTCGAGATCTCTGATGCGCGGCTAAT---CGAAAAGCTGG 1485
 Qy 1520 TCCAGGAGCATGCGACATTAATCTCTGTCAGACCTATGAGCATCAAGGATTTACT 1679
 Db 1486 TCGAAAAGCTGCCCCGATTTCTGTAACCTGCGGAAGCTTCGGCTGGAAGGGATCATC 1545
 Qy 1680 GTGCGCAAGCGGAAAGATTTGGCCCCGCGCATGCGGAAATGCTAGCCCAATGCTCT 1739
 Db 1546 TGCCGCGATCCGAGAGAGCTGAGCATGCGATGAGATGATCGACTGAGATGCTCG 1605
 Qy 1740 GTGTCATGATGTCGTCGTCGCAAAAAGATGAAAGTATCCCTATGATGCGCCCGG 1799
 Db 1606 GTGCTTTTGTATGTCGTCGTCGAGAGACGAGATGCTTCCGATGATCCGTCGCG 1665
 Qy 1800 ATGAGTATGTCGCAATGCTAGTTTACCGGAAGTGCCTGACGCAATGCTCCCGG 1859
 Db 1666 AAGCGCATACGAAATGCTCTCTCGGCGCATGCGGCGCTTTGCGGCGGCGCGCG 1725
 Qy 1860 ATGGTG 1865
 Db 1726 CTGGTG 1731
 RESULT 14
 AAS52288
 ID AAS52288 standard; DNA; 1815 BP.
 AC AAS52288;
 DT 13-FEB-2002 (first entry)
 DE E. coli DNA for cellular proliferation protein #10.
 KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 KW antibacterial; drug design.
 OS Escherichia coli.
 PN W0200170955-A2.
 XX 27-SEP-2001.
 PD 21-MAR-2001; 2001WO-US009180.
 PF 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI: 2001-611495/70.
 DR P-PSDB; ANU34429.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 PS Claim 27; SEQ ID NO 5925; 511pp; English.
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes

CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 1815 BP; 419 A; 424 C; 529 G; 443 T; 0 U; 0 Other;
 Query Match 18.8%; Score 368.6; DB 4; Length 1815;
 Best Local Similarity 55.3%; Pred. No. 6.5e-106;
 Matches 805; Conservative 0; Mismatches 634; Indels 16; Gaps 4;
 Qy 172 CATGGGTCMAACACATTTTGGCTATCCCGCGGGGCATTTTGGCCATGATGAA 231
 Db 139 CAGGGCGTTAAACAAGATTCGGTATCCCGAGGCGGACGCTGATATTATGATGCA 198
 Qy 232 CTGTACCGCTTGAACCGCGGGGAAATTGACATATTGTCGCCGCAACAGAGA 291
 Db 199 -----TTGCATACCGTGGTGGTATGATATGATGATGATGTCATGACACAGCG 249
 Qy 222 GCTTCCATGCGCGGATGGGATGCGACAGCCAGAGTAAGGAGATTGTTCCGT 351
 Db 250 GCGGTGATATGCGGATGCGCTGCGCGCCGACCGGGAAGTCGGTGTCTGTA 309
 Qy 352 ACATGTGACACGCGGCGCATTAATCTGTCACCGGATTCGCAATGCCATTGGACTG 411
 Db 310 ACGTGGGTCAGGGGCGACCAATGCGATTAATGTCATGCGACGACGCTTAATGATTC 369
 Qy 412 GTGCCCATGTCGTCATTAATGTCGAGAGGTCGCGTCGTCATGATGTCGTCCTTC 471
 Db 370 ATTCATTAATGTCCTTCCTCCGAGAGTACGACCTGTCGATAGTTAGATCCCTT 429
 Qy 472 CAGAAATTAATGATTTTGGATCCATCTTACCGATGTTAAACATCTTATGTCAGT 531
 Db 430 CAGGAGTGGACATGTCGGGATTTCCGCAACGGGTGTTAAACAGATTTCTGGTTAG 489
 Qy 532 AGTGGCGGAGATGTCGTCGATTTGTAATGAGGCTTCATCTGTCAGACCGGTCGT 591
 Db 490 CAACGGAAGACATTCGAGGTCGTAAGAAAGCTTTCTGCTGCGGCAAGTCGTC 549
 Qy 592 CCGGCGCGCTTTTATGATGATATTCGCAAGATGTCGCTTAAAGAAATGTCATCT 651
 Db 550 CAGAGACAGTATGTCGTTGATTCGAAAGATTTCTTAAATCCGGCAACAAATTACC 609
 Qy 652 CCCCTGACCCCGGTCGATTAATCTACCGGCTTACGCCCAAGGTTAAGTAACTCC 711
 Db 610 TATGTCGTCGAGTCGATGTCGATGTCGTTTAAATCCCATCTACCGACATAA 669
 Qy 712 CGACAAATTAATGCGGATTCGATGTCGATGTCGAGCGGCAAGATCCCTTGCTACGTA 771
 Db 670 GGGCAATTAATGACGTCCTTCGAAAGCTGTAAGGCGGCAAAAACCGGTTGTACGTA 729
 Qy 772 GGGGAGGGGCGATGCGCGCAATGCGCAATGCGGTCGAGAGAAATTTGCGGAAGTTC 831
 Db 730 GGGGTCGGGCAATACGCGCGGCTGCAATGCGAGTGAAGAAACGATGAGCGCTTG 789
 Qy 832 CAGTTCGCGTAAACAACCTGATGGAATTTGGGGCTTTTGAAGAAACCATCCCTT 891
 Db 790 AATCTGCGGTCGTTGTCATGTAAGGCTGCGGCGCTTTCCGCAAGCATGTCAG 849
 Qy 892 TCGTGGTATGTCGTCATGTCAGCCACCGCTATGCCAATTCCTGTCGCAAGTGT 951
 Db 850 GCACTGGATGTCGTCATGTCAGCGTACCTACGAAGCAATATGATGATGATACGCG 909

```
QY 952 GATTGTGATGTCAGTGGGAGCCGTTTCAGACGACCGGTAATCGCAAACTAGACGAA 1011
Db 910 GATGTATATTTTCGCCGCGGGGTACGATTTGATGACCGAAGCAAACTCTGGGAAG 969
QY 1012 TTGTCTAGCCGCCCAAGTAATTCATTGATGATGACCCCGGAGGTGGAAAAAAC 1071
Db 970 TACTGCCCAATTCGCACTGTTGCAATTCGATATGATTCCTCACTTCCTTAAACC 1029
QY 1072 AGGGCTCCGATGTGGCCATTGTGGGGATGTACGCCATGTTTATAGACGCTTTGACG 1131
Db 1030 GTACCTCGGATATCCGANTGTGGGGATGTCTCCGCAAGTCTCTGCAAAATGCTTGA 1089
QY 1132 CGGGCCCGGAATGATTAACCCACCCATCCCATACACCCAGGATGTTAAATGCG 1191
Db 1090 CTCTTGTGCAAAATCCGCCCATCAACCACTGGATGATCCGCACTGTGGACGAA 1149
QY 1192 ATGTATCATTTGGCGGACCCATTCACCCCTCCAGGTCCCACTATGAGATCTATGTC 1251
Db 1150 ATTGAACAGTGGCGCGCTGTCAGTCCCTGAAATATGACACTCAAGTAAAGATTAA 1209
QY 1252 CCCAGAGGTATGACGAAATTTGTCGCGAGGCC---CCGATGCTTATACACCAAC 1308
Db 1210 CCCAGGCGGTGATCGAGACTTTTGGCGTTGACGAGGAGAGACGCTTACGTGACGTC 1269
QY 1309 GATGTGGGCAACACCAAAATGTGGGGGCGCC--AGTTTTGAACAATGGCCCCCGCGA 1365
Db 1270 GATGTGGGAGACCAAGATGTTGCTGCACTTATATCACTTGAACAAACCGCGTGC 1329
QY 1366 TGGATTTCCAGTGTCTGTGGGTAGCATGGCTTGTGTTACTGCGCCCATGGAGCC 1425
Db 1330 TGGATCAATTCGGGTGCTCGGCGCATGAGGCTTTGTTACTGCGCATGCGCGCTC 1389
QY 1426 AAGTGGGAGTGGGGGACGAG--CGTCAATTTGATGATGAGATGCGACGCTTCAATG 1484
Db 1390 AAAATGGCGTTGCAAAAGAAACCGGTGTTTGGCTGACCTGGCGAGCGAGTATCAGATG 1449
QY 1485 AATCTTCAGGAATGAGAACCTTAGCCAGTACGACATCCAGGTTAAACTATTTCTC 1544
Db 1450 AATCTTCAGGAATGATGCTACCCGCTTGCAATCGAGTGGCCCTACTGCTGATCTC 1509
QY 1545 AATTAAGTGTGGAGGGATGTGTGCTGATGAGTGGCAAACTTTTACGAGAACGTTAT 1604
Db 1510 AATTAAGTGTGTGAGGATGTGTGAGTGAAGCATGGGAGGACATGATCTATTCGCGCTCAT 1569
QY 1605 TCGGCTTACATG 1619
Db 1570 TCAATCTTATATG 1584
```

```
RESULT 15
ACA32352
ID ACA32352 standard; DNA; 1815 BP.
AC ACA32352;
XX
XX
XX 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #14009.
XX
XX Antisense; ds: prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Escherichia coli.
XX
XX WO20027183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002MO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
PR
```

```
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX P-PsDB; ABU28482.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 20222; 1766p; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1815 BP; 419 A; 424 C; 529 G; 443 T; 0 U; 0 Other;
XX
XX Query Match 18.8%; Score 368.6; DB 7; Length 1815;
XX Best Local Similarity 55.3%; Pred. No. 6.5e-106;
XX Matches 805; Conservative 0; Mismatches 634; Indels 16; Gaps 4;
```

```
QY 172 CATGGGGTCAACAACTTTTGGCTATCCCGGGGGGCAATTTGGCCATCTATGATGA 231
Db 139 CAGGGCGTTAAACAAGATTCGTTATCCCGAGGGCGAGTCTTATATTTATGATGA 198
QY 232 CTGTACCGCTTGAAGGGCGGGGAATTAACATATTTTGGTGGCCATGAACAGGA 291
Db 199 -----TTGCATACCGTGGGTGGATGATGATGATATATGATATGACAGGCG 249
QY 292 GCTTCCATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351
Db 250 GCGGTGATATGCGCGATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTA 309
QY 352 ACATCTGACCAAGGGCGCACTAATGATGACCGGATGATGATGATGATGATGATGATG 411
Db 310 ACGTGGGTGCAAGGGCGCAATGATGATGATGATGATGATGATGATGATGATGATG 369
QY 412 GTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 471
```

```
Db      | 370 | ATTCATTAAGTTGTCCTTTCCGGCAGGTAGCGACTCGTTGATAGGTTACGATGCTTTT 429
Qy      | 472 | CAGGAATTTGACATTTTGGCATCACCTTACCGATTCGTTAAGCACTCTTAAGGTACGT 531
Db      | 430 | CAGGATCGCAATGATGGGATTTCCGACCGGTGTTAACAAGTTTCTGATTAG 489
Qy      | 532 | AGTGGCGGATATGCTCGCATTTGTTACTGAGGCTTTCCATCTTGTACACCGGTCT 591
Db      | 490 | CAACGGAAAGACATTCGCGCAGGTCTGAAAAAGGCTTCTGTGCGCAAGTGTGCG 549
Qy      | 592 | CCGGGCGGTTTGTATGATGATATTCACAGGATGCGGCTTAGAAGATGTGATATT 651
Db      | 550 | CCAGACCAATGATGCTGATTTACCGAAGATATTTCTTAATCCGGGAAACAAATTACC 609
Qy      | 652 | CCCCCTGACCCCGGTGACGTTAATCTACCGGTTATGCCCCACGTTAAAGTTATCCC 711
Db      | 610 | TATGCTGCGCGAGTGGTCAGTATGCTTCTTAACAATCCACTACTACCGGACATPAA 669
Qy      | 712 | CGACAAATTAATGCGGCAATTGCAATGTTGAGACAGCCAGAAATCCCTTGCTTACGTA 771
Db      | 670 | GGGCAAAATTAAGCGTGTCTGCAAACTGTGTAGCGGCAAAAAACCGTTGTCTACGTA 729
Qy      | 772 | GGGGAGGGGCGATGCGCGCAATGCCCAGTCCCGAGTGCAGAAATTTGCGAAAGGTT 831
Db      | 730 | GCGGATGGGCAATACCGCGGCGTCCATCAGCAATTGAAGAAACGTGAGGCGTTG 789
Qy      | 832 | CAGTTCCGCTAACCAACCTGTATGGAATTTGGGCTTTTGAAGAAACATCCCTT 891
Db      | 790 | AATCGCCGTTGTTGCTCATTTGATGGGCTGGGGGCTTTCCGGCAACGATCGTCAG 849
Qy      | 892 | TCGGAGGATATGTTGGGTATGATGATGCAACCGCTATGCCAACTTGCCGTACGGAATG 951
Db      | 850 | GCATGGGCAATGCTGGGAATGCAAGTACTTAAGAACCAATATGACATGATCAATACGG 909
Qy      | 952 | GATTGTTGATTTGAGTGGGCGCCCGTTTCAGACGACCGGTAATGCGAACTAGACGAA 1011
Db      | 910 | GATGTGATTTTGCCTGGGGTACGATTTGATGACCGAACGAAACAACTGCGCAAG 969
Qy      | 1012 | TTTGCTACCGCGCAAGTAAATTCACATGACATGCAATGCAACCGGCGGAGGTGGA AAAAC 1071
Db      | 970 | TACTGCCAATATGCACTGTCTGCAATGCAATGATGATGCTTCTCATTTCTTAAC 1029
Qy      | 1072 | AGGCTCCCGATATGTCCTCATTTGTGGGAGTGTACGCCATGTTTGAACAAGCTTTGCA 1131
Db      | 1030 | GTGATCTGGGAATATCCGATTTGTGGGAGTGTGCGCAGTCTCGAAACAAATGCTTGA 1089
Qy      | 1132 | CGGCGCCGGAATTTGATTTACCCCATCCATCCCATPACACCGCAGGCAATGTTAATG 1191
Db      | 1090 | CTCTTGCGCAAGAAATCGCCCATCAACCACTGATGATCCGACTGTGCGACGAA 1149
Qy      | 1192 | ATTGATCATTTGCGGACCGATTAACCCCTCCAGGTGCCCCACTATGAGATATATTGCC 1251
Db      | 1150 | ATTGAACATGCGCGCTGCTGCAATGCTTGAATATGACACACAGTGAAGAAATTA 1209
Qy      | 1252 | CCCCAAGAGTATGACAGAAATTTGTGCGCAAGGCC--CGATGCTTACTACACCA 1308
Db      | 1210 | CCGCAGGCGGTATGAGACTCTTTGGCGGTTGAGAAAGGAGACGCTTACGTGACGTCC 1269
Qy      | 1309 | GATGTGGGACACACCAATATGTTGGCGGCC--AGTTTGAACAATGCCCCCGCGCA 1365
Db      | 1270 | GATGTGGGACACACCAATGTTTGTGCACTTTATTCATTTGACAAACCGGTGCG 1329
Qy      | 1366 | TGGATTTTCAGTGTGCTTGGGTAAGATGGGCTTTGTTTACTGCGGCAATGGAAGCC 1425
Db      | 1330 | TGGATCAATTCGGTGGCTCGGACAGATGGGTTTGGTTTACCTGCGGCACTGGGCGTC 1389
Qy      | 1426 | AAGTGGAGTGGGGAGACAG--CGTCAATTTGCAATCAGTGAAGTCCAGCTTCCAATG 1484
Db      | 1390 | AAAATGGCTTCCAGAAAGAAACCGTGTGCTGCACTGGGAGCGCATTTCAATG 1449
Qy      | 1485 | AATCTTCAGGAATGAGAAACCTTACCGCATGATCAGGTTAAATATTATTCTC 1544
```

```
Db      | 1450 | AACATCCAGAACTGTCTACCGCGTTGCAATACGATGCCCGGACTGTGATGATCTC 1509
Qy      | 1545 | AATAACGTTTGGCAGGGATGTGCGTCAGTGGCAACAACTTCTACGAAGAACGTTAT 1604
Db      | 1510 | AATAACGCTATCTGGGATGTGTAAGCAGTGGCAGACATGATCTATTCGGGCGGTAT 1569
Qy      | 1605 | TCTGCTTCAATG 1619
Db      | 1570 | TCACATCTTATATG 1584
```

Search completed: July 25, 2004, 05:27:58
Job time : 830.718 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 05:10:12 ; Search time 152.901 Seconds
(without alignments)
7110.132 Million cell updates/sec

Title: US-09-893-033-6
Perfect score: 1959
Sequence: 1 gccatagagcccatgcgcg.....cataagcccaaatgaattc 1959

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/pdata/2/ina/5A.COMB.seq: *
2: /cgn2_6/pdata/2/ina/5B.COMB.seq: *
3: /cgn2_6/pdata/2/ina/6A.COMB.seq: *
4: /cgn2_6/pdata/2/ina/6B.COMB.seq: *
5: /cgn2_6/pdata/2/ina/PCTUS.COMB.seq: *
6: /cgn2_6/pdata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	416.6	21.3	1740	4	US-09-489-039A-5163 Sequence 5163, Ap
2	370.6	18.9	1716	4	US-09-252-991A-7372 Sequence 7372, Ap
3	370.6	18.9	1743	4	US-09-252-991A-7128 Sequence 7128, Ap
4	368.2	18.8	96109	4	US-09-596-002-35 Sequence 35, Appl
5	366.8	18.7	4411529	3	US-09-103-840A-1 Sequence 1, Appl
6	365.4	18.7	4403765	3	US-09-103-840A-2 Sequence 2, Appl
7	353	18.0	1632	4	US-09-540-236-646 Sequence 646, App
8	328.4	16.8	1830121	4	US-09-557-884-1 Sequence 1, Appl
9	328.4	16.8	1830121	4	US-09-643-990A-1 Sequence 840, App
10	323.8	16.5	1689	4	US-09-489-039A-840 Sequence 840, App
11	321.6	16.4	1731	4	US-09-328-352-3942 Sequence 3942, Ap
12	318.8	16.3	2111	4	US-10-096-571-11 Sequence 11, Appl
13	318.8	16.3	2111	4	US-10-096-571-13 Sequence 13, Appl
14	309.2	15.8	2841	4	US-08-452-075-1 Sequence 1, Appl
15	309.2	15.8	2841	3	US-09-231-061-1 Sequence 1, Appl
16	309.2	15.8	2841	3	US-09-231-061-1 Sequence 5, Appl
17	304.4	15.5	1947	4	US-09-489-039A-3367 Sequence 3367, Ap
18	304	15.5	6211	4	US-08-961-527-8 Sequence 8, Appl
19	293	15.0	1969	1	US-07-737-851-1 Sequence 1, Appl
20	293	15.0	1969	1	US-07-737-851-2 Sequence 2, Appl
21	293	15.0	1969	1	US-07-894-062-1 Sequence 1, Appl
22	293	15.0	1969	1	US-07-894-062-2 Sequence 2, Appl
23	293	15.0	1969	4	US-09-096-562-1 Sequence 1, Appl
24	293	15.0	1969	4	US-09-096-562-2 Sequence 2, Appl
25	290.6	14.8	1728	1	US-08-403-866-12 Sequence 12, Appl
26	290.6	14.8	2231	1	US-08-403-866-14 Sequence 14, Appl
27	290.6	14.8	12720	1	US-08-403-866-11 Sequence 11, Appl

28	288.2	14.7	1969	1	US-07-894-062-3 Sequence 3, Appl
29	288.2	14.7	1969	4	US-09-096-562-3 Sequence 3, Appl
30	286.6	14.6	1969	1	US-07-737-851-3 Sequence 3, Appl
31	276.2	14.1	1020	4	US-09-252-991A-7248 Sequence 7248, Ap
32	275	14.0	1664976	4	US-08-916-421B-1 Sequence 1, Appl
33	271.8	13.9	640681	4	US-09-790-988-1 Sequence 1594, Ap
34	271.6	13.9	1710	4	US-09-543-681A-1594 Sequence 1, Appl
35	269.4	13.8	2156	1	US-08-321-356-1 Sequence 3, Appl
36	269.4	13.8	2156	1	US-08-321-356-3 Sequence 1047, Ap
37	256.6	13.1	1755	4	US-09-134-001C-1047 Sequence 199, App
38	254	13.0	1689	4	US-09-543-681A-199 Sequence 1, Appl
39	240.8	12.3	2365	3	US-08-363-208-1 Sequence 102, App
40	240.8	12.3	2365	3	US-09-137-478-1 Sequence 102, App
41	230.8	11.8	15249	4	US-08-956-171E-102 Sequence 11016, A
42	155.4	7.9	1869	4	US-09-252-991A-11016 Sequence 10943, A
43	155.4	7.9	1864	4	US-09-252-991A-10943 Sequence 11240, A
44	155.4	7.9	1959	4	US-09-252-991A-11240 Sequence 881, App
45	130	6.6	315	4	US-09-489-039A-881

ALIGNMENTS

RESULT 1									
US-09-489-039A-5163									
Sequence 5163, Application US/09489039A									
Patent No. 6610836									
GENERAL INFORMATION:									
APPLICANT: Gary Breton et. al									
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA									
TITLE OR INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS									
FILE REFERENCE: 2709.2004001									
CURRENT APPLICATION NUMBER: US/09/489,039A									
PRIOR FILING DATE: 2000-01-27									
PRIOR APPLICATION NUMBER: US 60/117,747									
NUMBER OF SEQ ID NOS: 14342									
SEQ ID NO 5163									
LENGTH: 1740									
TYPE: DNA									
ORGANISM: Klebsiella pneumoniae									
US-09-489-039A-5163									
Query Match									
Best Local Similarity 21.3%; Score 416.6; DB 4; Length 1740;									
Best Local Similarity 55.8%; Pred. No. 8.8e+128;									
Matches 905; Conservative 0; Mismatches 699; Indels 19; Gaps 5;									
QY	172	CATGGGGTCAACACATTTTGGCTATCCGGGGGCAATTTGGCCATCTATGATGAA	231						
DB	64	CAGGGCTCAAGCAAGATTCGGCTATCCGGAGGGCGAGTCTCGATATCTATGATCG	123						
QY	232	CTGTACCGCTTTGAGCGCGGGGAAATTGACATATTTGGTGGCCATGAACAAGA	291						
DB	124	TTACATACCTTTG-----GCGGTATCGATCATGTGTGTTGTCACGACAGCG	174						
QY	292	GCTTCCATCGCGGGATGATGCAAGCCAGAGGTAAATGGAGATTGTTCCGT	351						
DB	175	GCGGTGATATGCGGATGCTCGGGGCGACCCGGGAGTCCGCTGTGCTGTG	234						
QY	352	ACATCTGAGACCAAGGGGAGTAACTTGTGACCGGATTTGCCAATGCCATTTGA	411						
DB	235	ACCTTGGGGCTGGGGGAGCAACGACATCATGGAATTCGACACCTATATGACTC	294						
QY	412	GTCGCATGATGATGATTTCTGAGAGTGGCGCGTCCATGATTTGATGATGCTTC	471						
DB	295	ATTTCGTTGATTCCTCTCGGGGAGGTAGCCACTCGCTTATCGCTACGATGCTTC	354						
QY	472	CAGGAATTTGACATTTTGGCATCACTTACCGATGTTAAGACATCTATGTGATCGT	531						
DB	355	CAGAGTGCACATGTGGGATTTCCCGTCCGTGTGAGACAGTTTCCTGTGAAG	414						
QY	532	AGTGGCGGATATGCTGCTGATTTTCTGAGGCTTTCCATTTGCTAGACCGGTCTG	591						

415 CAGACGAGATATTTCCCGGCTGCTGAGAAAGCCTTCTGCTGCGGCGCAGCGCTGC 474
 QY 592 CCCGCGCGCTTTTGTATGATATTTCCAAAGATGTGGGCTTTAGAAAGATGTAGTACAT 651
 Db 475 CCCGCTCGGATGATGATCTGCTGAGAGATATTTCTCAATCCGCGGAAAGAGCTGCT 534
 QY 652 CCCCTGACCCCGGATGATATCTACCGGCTTATGCGCCCAAGGTTAAAGGTAATCC 711
 Db 535 TACGCTGCGGAGCGAGTGAAGTGCCTGCTGTAACCCCAACACCAAGTGCATATAA 594
 QY 712 CGACAAATTAATGCGGATTCGCAATTTGTGAGACAGGCGAAGAAATCCCTGCTTACGTA 771
 Db 595 GGGGAGATTTAAACGGGCGCTACAGACTGTGTGGCCGAGTAAGTCCGCTGCTTACGTC 654
 QY 772 GGGGAGGCGGATGCGCGCAATGCCCAGATGCCAGATGAGGAAATTTGGGAAAGTTG 831
 Db 655 GCGCGCGGCGGATCAATGCGATTCGCAACCTTACGTCGCGGCTGTGGGAAATCTG 714
 QY 832 CAGTTGCGGATTAACAACACCCCTGATGGGAAATTTGGGCTTTTACGAAATCCCTT 891
 Db 715 AAGCTGCGGCTGCTCATCTTGTATGAGACTTTGGCGCTTTCCGCGGAGACACTGCGAG 774
 QY 892 TCGGCTGATGTTGGGATGCAATGCGCACCGCTTATGCGAACTTTGCGCTGAGGAAAT 951
 Db 775 GCGGCTGGGATGCTGGGATGCAAGGCACTTATGAGCAATGACTATGCACTGCG 834
 QY 952 GATTGTTGATGCAATGCGGCGGCTTTCAGACAGCGGCTTACGCTGCAACTAGACGA 1011
 Db 835 GACGTGATTTTCCGCTGCGCTTCTGCTTGAACACCGGAGACCACTTACCTGCGCAAG 894
 QY 1012 TTTGCTAGCCGCGCAAGTAATTCACATTGACATTCGACCGCGGAGGTGGGAAATAC 1071
 Db 895 TATTGTCCGAACGCGACGCTACTGATGATGATGATGATGATGATGATGATGATGAT 954
 QY 1072 AGGGCTCCCGAATGCGGCTTGTGGGAGTATGAGCACTTTTGAACAGCTTTTGCAG 1131
 Db 955 GTGCGGAGATGTCGATGCTGCGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 1014
 QY 1132 CGGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
 Db 1015 CTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1074
 QY 1192 ATTGATTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1251
 Db 1075 ATCGAGAGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
 QY 1252 CCCGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
 Db 1135 CCGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 QY 1309 GATGTTGAGCAACACCAATGTGGGCGGCGCAATTTT---TGAAATGAGGCGGCGCA 1365
 Db 1195 GATGTTGAGCAACACCAATGTGGGCGGCGCAATTTT---TGAAATGAGGCGGCGCA 1354
 QY 1366 TGGATTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1425
 Db 1255 TGGATTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314
 QY 1426 AAAAGTGGAGTGGGAGGAG-AGCGGTCAATTTGATCAGTGGAGATGCGAGCTTCAATG 1484
 Db 1315 AAAAGTGGAGTGGGAGGAG-AGCGGTCAATTTGATCAGTGGAGATGCGAGCTTCAATG 1374
 QY 1485 AATCTTCAGGAACTGCGGAACTTACGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1544
 Db 1375 AATCTTCAGGAACTTACGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1434
 QY 1545 AATAACGATGAGCAGGAGATGATGCTGATGAGCAACAACTTTCTACGAAAGACCTTAT 1604
 Db 1435 AATAACGATGAGCAGGAGATGATGCTGATGAGCAACAACTTTCTACGAAAGACCTTAT 1494
 QY 1605 TCTGCTTTTAACTATGCTGAGGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1664
 Db 1495 TCGAGTCTTATATG---GAAATGCTACCGGACTTTTGTGCTGCGGAGGAGGCTTATGCT 1551

QY 1665 ATCAAGGATATTAATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1724
 Db 1552 CAGCTGGGATTCGCGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1611
 QY 1725 GCCCAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
 Db 1612 GAGCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1671
 QY 1785 ATG 1787
 Db 1672 GTG 1674

RESULT 2
 US-09-252-991A-7372/c
 ; Sequence 7372; Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubinfeld et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 7372
 ; LENGTH: 1716
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-7372

Query Match 18.9%; Score 370.6; DB 4; Length 1716;
 Best Local Similarity 54.5%; Pred. No. 1.8e-112;
 Matches 884; Conservative 0; Mismatches 714; Indels 25; Gaps 6;

QY 135 AACCGGCGCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 194
 Db 1714 ATCTGGCGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
 QY 195 CTATCCCGGCGGCGCAATTTTGGCATCTGATGATGATGATGATGATGATGATGATGATGAT 254
 Db 1654 GTACCGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1604
 QY 255 GGAATGAGCATATTTTGGTGGCGCATGAAACAAGAGCTTCCATGCGGCGGAGTGA 314
 Db 1603 CAGCTGACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1544
 QY 315 TCCCAAGGCAAGTAAAGTGGAGTTTGTTCGTAATCTGTAACCAAGGCGAGCTTA 374
 Db 1543 CCGCGCGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 1484
 QY 375 CTGGTGAACGCGCATTTGCAATGCGCATTTGCAATGCGCATTTGCAATGCGCATTTGCAAT 434
 Db 1483 CCGCATACCGGTATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 1424
 QY 435 AAGAGTGGGCGGTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
 Db 1423 CAGAGTGGGCGGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1364
 QY 495 CACCTTACCGATGCTTGAACATCTCTATGCTGATGATGATGATGATGATGATGATGATGAT 554
 Db 1363 CTCCCGGCGGATGGAAGCAAGCTTATCATCATCAAGATCCCTTGGAAATTCGAGAGT 1304
 QY 555 TGTACTAGAGCTTTCATCTTGTATGACACCGGCTGCTGCGGCGGCTTTTGTATGATGATAT 614
 Db 1303 GATCAAGAGGCGCTTTTACCTGCGGAGTCCGCTGCTGCGGCGGCGGCTGCTGCTGCTGCT 1244
 QY 615 TCCCAAGGATGTTGGCTTGAAGAAATGTAAGTATTCCTCCCTGACCCCGGAGAGCTTAA 674

Db 1243 TCCGAGGACATGGGCGACCGACGACGAGTTCGATATTCCTATCCGAAAGAGTCA 1184
QY 675 TCTACCGGGTTATCGCCCAAGGTTAAAGGTATCCCGACAAATTAATGCGGATTGCA 734
Db 1183 GTTGCGTTCGTAACACCCCGCGGTTCGCGGTCACTCGGACAGATCCCGAAGGCGCCGA 1124
QY 735 ATTGTTGAGCAGCGCAGAAATCCCTGTCTAGCTAGGAGGAGGCGGATCGCGCCAA 794
Db 1123 GATGTCCTCGGCGCCGACAGCGCCCGGTGTCTATTCGCGCGCGCGGTGATCATAGGCA 1064
QY 795 TGCCCATGCCAGGTGACGAAATTTGCGAAAGTTCCAGTTGCGGTPACACACCCCT 854
Db 1063 TGCCCGCGCGCGGTGACCGAGCTGGCGCAGATGCTCACTGCGCGGTGACCAACACCT 1004
QY 855 GATGGAAATTGGGGCTTTTGACGAAACCATCCCTTCGGTGGGTATGTTGGTATGCA 914
Db 1003 GATGGGCTCGCGGGTATCCCGCGCGACCGCATCTTCGTCATGCTCGGATGCA 944
QY 915 TGGCCACCGCTATGCGCACTTTGCGGTGACGGAATGATTTGATTGAGTGGGCGC 974
Db 943 CGGACGCTTCAACCGCAACCTGGCGATGATCAACGACGATGATCCGCGGTGCGCGC 864
QY 975 CGGTTTCAACGACCGGGTATCTGGCGAACTAGACGAATTTGCTAGCCCGCGCAAGTAT 1034
Db 883 GCGTTTCAACGACCGGGTATCTAGACGCGCGCGCAAGTTGCGCGCAACGCAAGTAT 824
QY 1035 TCACATTGACATCGACCGCGCGGAGGTGGGAAACAGGGGTCCCGCATGTCGTCATTTG 1094
Db 823 CCACATCGACATCGATCCGCGGTGATTTCCAAAGCATCAAGGCGCGCATCCGATGT 764
QY 1095 GGGGATGTACCGCATGTTTGAACAGCTTTTGACGCGGCGCGGAAATGATTAACC 1154
Db 763 CGGCGCGGTGACGACGTTTTCACCGAAATGGTCGATCGTCAAGGAAATCGCGAGAC 704
QY 1155 CACCATCCCATPACACCCAGGCGATGTTAAATGCGATTCATTAATGCGC---GACCGCA 1211
Db 703 CCGCAACGAGATGCCAGCGCTGCTGTGTGAGAGATCAACGAGTGGTGTATACG 644
QY 1212 TTACCCCTCCAGTGGCCCATATAGAGATTAATTCGCCCCCGAGAGAGTATGACGA 1271
Db 643 TGGGTTGTCCTGTCGACAGAGGCGACGCGACATCATCAAGCGCGACGCGTATGCA 584
QY 1272 AATTGTCGCGCGCC---CGATGCTTACTACACACCGATGTTGGACACACCA 1325
Db 583 GACGCTTACGAGGTTTACCGATGCGATGCTTATCATCCTTCGATGTCGCGACGACCA 524
QY 1326 AATGTCGCGCGCGCGATTTTGA---CAATGCGCGCGCGATGATTCAGTGCAG 1382
Db 523 GATGTCCTGTCGACGATCAAGTTCACAGCCCATGCTGATCACTCGCGCG 464
QY 1383 CTGGGATCGATGGCTTTGTTTACCTGCGCCATGCGAGCGAAAGTGGAGTGGGGA 1442
Db 463 TCTCGGACACATGGGCTTCGCTCCGCGCGCATGGCGATCAAGCTCACTTCCGGA 404
QY 1443 CGA-GCGGCTTTCATCATGTCGAGATGCGAGCTTCCAAATGATCTTCAGAAATCGG 1501
Db 403 CGACATGTCCTGTCGTCGACGCGAGGCGATTCAGATGAATTCAGAGAACTGTC 344
QY 1502 AACCTGACCCGATGACATTCAGAGTTAAACTATTAATTCATATAACGTTGGACAGG 1561
Db 343 TACCTGCTGACGATCGACCTGCGGTGAAGTCTCACTCAACATGATGCGCTGAG 284
QY 1562 GATGTCGTCAGTGGCAACAACTTTCTAGAGAAAGTTATTCGCTTCAACATGTC 1621
Db 283 CATGTCGCGCAGTGGAGAGATGACATGACAGCGGTTATTCGATTCCTCATG- 226
QY 1622 CCAGGACATGCGACATTAATTCCTCTGTCGAGCTTATGAGCATCAAGGATATTAAT 1681
Db 225 -GAATCCCTGCGGACTTGTCAAGCTTGCAGAACTCAACGAGCATGTCGCGAT 167
QY 1682 GCGCAGCGGAAATTTGGCGCGCGGATGCGCGAAATGCTAGCCCAATGTCCTGT 1741

Db 166 CACGACCTGAGAGATCTGAGCCGGAAGATGAGAGAGGGTTGCGCATGAGATGCGCT 107
QY 1742 GGT 1744
Db 106 GGT 104

RESULT 3
US-09-252-991A-7128
; Sequence 7128, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7128
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7128

Query Match 18.9%; Score 370.6; DB 4; Length 1743;
Best Local Similarity 54.5%; Pred. No. 1,8e-112;
Matches 884; Conservative 0; Mismatches 714; Indels 25; Gaps 6;

QY 135 AACCGGGCTTATATCTGATGATGATGAGTAAACGCCATGGGGTCMAACACTTTTGG 194
Db 30 ATGTGGCGCTGAATAGTGTGTCGCTGTCGCGCGCGAGAGGCTTAAGTACATCTACG 89
QY 195 CTATCCCGGGGCGCAATTTTGGCCATCTATGATGAATGATCGCTTGAACGGCGG 254
Db 90 GTACCGGGGTGGTCCCTCCCTGATATCTACGACGCGCTCTT-----AAGAGGA 140
QY 255 GGAATTGAGCATTTTGTGTGCGCATGAACAAGAGTTCATCGCGCGATGGGTA 314
Db 141 CGAGTACCCACATCTGTCGTCGCGCAACAGCGGCTTACCCATATGACGACGCTA 200
QY 315 TGCCAGGCGCACAGTAAAGTGGAGTTTGTTCGTCATCTGACATCGACAGGCGCAT 374
Db 201 CGCCCGGCGCACCGGACAGCGCGGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 260
QY 375 CTGTGTACCGCGCATTTGCCATATGCCATTTGACTGCTGTCGTCGTCGTCGTCGTC 434
Db 261 CGCATATCACCGGATGACGACCGCATATGACATGACATGACATGACATGACATGAC 320
QY 435 AGAGTGGGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 494
Db 321 CCAGTTCGCGACGACATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 380
QY 495 CACCTTACCGATGCTTAAGCATCTCTATGTGTGTCGTAAGTGGCGGATGATGCTGCAT 554
Db 381 CTCGCGCGCGATGTCGACGACAGCTTCACTCAACGATTCCTCCGAATTCGCCAGGT 440
QY 555 TGTATGAGGCTTTCATCTTCTGACACCGGTCGTCGTCGTCGTCGTCGTCGTCGTCG 614
Db 441 GATCAGAAAGGCTTTTACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 500
QY 615 TCCCAAGATGAGGCTTGAAGAAATGATGATGATGATGATGATGATGATGATGATG 674
Db 501 TCCGAAGACATGGGCGACCCGACGACGAGAAATTCGATATTCCTATCCGAAGAGTCA 560
QY 675 TCAACCGGTTATCGGCGCACCGGTTAAGGTATCCCGCAATTAATGCGGCATTGCA 734
Db 561 GTTGCTGTACAGCCGCGCGCTTCGCGTCACTGGAGACGATCCGAGAGCGCCGA 620

[illegible]

```
QY 797 CCCATGCCAGGTGACAGATTTGGCGAAGGTTCCAGTTGCCGGTAAACACCACTCTGA 856
Db 3084 CCAGTAAAGAGCTGGCTCAATTAGCCGATACGCTCAACTACTCTTACTAATACCTCTTA 3143
QY 857 TGGGAATTTGGGCTTTTACGAAAAACATCCCTTTGGTGGATATGTTGGATGATGATG 916
Db 3144 TGGGACTGGGCTCTTATCCAGGCTCAGCGGAACTTTTGGATGCTGGTATGATGATG 3203
QY 917 GCCACGCTATGCAACTTTGCCGTACAGGATGATTTGTTATTCAGAGTGGGGGCC 976
Db 3204 GCACATATGAGCGAATCATGACCAATGATCATGCTGATGATGATTTGGCAGTGGGTGTC 3263
QY 977 GTTTCGACGACCGGCTTAATGCGCAAACTAGACGAATTTGCTAGCCGCCCAAGTATTC 1036
Db 3264 GTTTTATGACCGGTGACCAATATATGCAAAAAATTTCCGCCAATGCAACGATCATTC 3323
QY 1037 ACATTGACATGACCCCGCGAGGTGGGAAAAAAGAGGCTCCGATGCTCCCATTTGTC 1096
Db 3324 ATATTGATATGATCCAGCGAGTATCTCAAGACGATTAAGCACACATTCGATGTTG 3383
QY 1097 GGGATGTAAGCCATGTTTATAGACAGCTTTTGACGGGCGCGGGAATTGGATTCCTCA 1156
Db 3384 GCGATGTAAGCTGTGACTAAATGAAATGCTTGCTTGCTGACAGAGATGATTAAGTGA 3443
QY 1157 CCCATCCCCATACCAACCGCATGTTAATGCAATGATCATGATTCGCGAC----- 1208
Db 3444 TCGATGAACATGCGCTCATGATTTGTGGCAACAAATCAATGATGCGTAAACGCCATG 3503
QY 1209 -----CGATTACCCCTCCAGGTGCGCCCACTATAGAGATATATGCCCCCAGAGG 1261
Db 3504 GTCTTCGCTATATGACATGATATGATGCTCCAAACACATGCAACGCCCAAGCGATG 3563
QY 1262 TAGTACAGAAATTTGTCGCCAGG---CCCCGATGCTTATACCAACCGCATGTCGAC 1318
Db 3564 TGGTATGAGTGTCTTATATAGCTGACGAATGGTGAAGCAATATACCTCCGATGTGCTC 3623
QY 1319 AACACCAATGTTGGGCGGCCAGTTT---TGAACATGGCCCCCGCGCATGATTCGA 1375
Db 3624 AGCATCAAGTGTCTGATTTGATATTAAGTATGATGACGACGCAATGTTAACT 3683
QY 1376 GTGTGCGCTTGGGTAAGTGGGCTTTGTTTACCTGCGGCATGGGACCAAGTGGAG 1435
Db 3684 CAGGTGGCTTAGGTACATAGGCGCTGGTTTACCTTATGCTATGCGACAAATTTAGCTC 3743
QY 1436 TGGGGGACGAG-CGGTCAATTGTCATCATGAGATGCCAGCTTCCAAATGATCTTACG 1494
Db 3744 ACCCAAAAAAGACATGTTGTTGATCACGTGGGAGGCTCAATCAATGAATCCAG 3803
QY 1495 AACTGGGAACCTTACCCAGTACGACATCCAGGTTAAACTATTCTCAATAAGCTT 1554
Db 3804 AGCTGTCACTTGTCTACAGTATATCTACAGTTAAATTTTAATTTAAATTAATGCTC 3863
QY 1555 GGCAGGGAGTGTGCTGATGTCAGCAAACTTTCTACGAAGAACTTATTTCTCTCTA 1614
Db 3864 AGCTTGGCATGTTAGAGCAGTGGCAGATATGCTTATGAGAGTGCAGCATCTCAATCTT 3923
QY 1615 ACATGTCCAGGGGATGCGACAGATTAATCTCTCTGTGAAGCCATGAGCATCAAGGTA 1674
Db 3924 ATATG---CAATCTTACTCTGATTTTGTAAATTTGGCAGAAAGTTATGGGATTAAGGTG 3980
QY 1675 TTACTGTGCGAAGCGGGAAGATTTGGCCCGCGATGCCGAAATGCTAGCCCAATG 1734
Db 3981 TAAATATTAACAATCTCTCTACATGACGAAGAGCTTAAACAGGCCCTTAAGATGATG 4040
QY 1735 GTCCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1794
Db 4041 GTTGGTATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4100
QY 1795 CCGCG 1799
Db 4101 CTGGC 4105
```

```
RESULT 5
US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 18.7%; Score 366.8; DB 3; Length 4411529;
Best Local Similarity 53.3%; Pred. No. 7e-109;
Matches 944; Conservative 0; Mismatches 772; Indels 34; Gaps 8;

QY 88 CCTAATCCGAACAAATTTCTGAAACTGTTCTTACGCAACGCAACCGGGCTTAT 147
Db 3362917 CCAAGACCCGGCGGCTGCTCGGAACATGTTGCACTGAGAGGTTACCGGTGACAG 3362858
QY 148 ATCCTGATGATAGCTGTGAACCGCATGGGGTCAAAACATTTTGGCTATCCCGCGGG 207
Db 3362857 GCGGTATCGGTCGTGGAGGAACTCGGGTGAAGTATTTCCGGATTCGGGGCGGT 3362798
QY 208 GCATTTTGGCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
Db 3362797 GCGGTGTGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362747
QY 268 ATTTGATGCGCATTAACAGAGAGCTTCCATGCGCGGATGGATGATGATGATGATGAT 327
Db 3362746 GTGCTGTGCGCATTAACAGAGAGAGCTTCCATGCGCGGATGGATGATGATGATGAT 3362687
QY 328 GTTAAAGTGGAGTGTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 387
Db 3362686 GGCCTGGTGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362627
QY 388 ATGCAATGCGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
Db 3362626 CTGACGACGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362567
QY 448 GCGATATGTTAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
Db 3362566 GGGCTATGCGGACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3362507
QY 508 GTTAAACATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
Db 3362506 ACCAAGCACATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3362447
QY 568 TTCCATCTTGTAGCAACCGGCTGCTCCGGGCGGTTTATGATGATGATGATGATGATG 627
Db 3362446 TTCCATCTTGTAGCAACCGGCTGCTCCGGGCGGTTTATGATGATGATGATGATG 3362387
QY 628 GCGTTAGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Db 3362386 CT-----GCAAGGCGCATGATGATGATGATGATGATGATGATGATGATGATG 3362333
QY 688 CGCCCGACGTTTAAAGTAAATCCCGGACAAATTAATGCGGATGATGATGATGATGATG 747
Db 3362332 AAGCCCAACACCAACCGGACCGGAGGTCGCGAGGCGCGGACGATGATGATGATG 3362273
QY 748 GCCAGAAATCCTTGTCTTACGATGAGGGGAGGCGGCGATGCCGCAATGCCATGCCAG 807
```

Db	3362272	GC	GGCGCAAGCGCGTGTGTATGTGGGGGGGGGCTCATCCGCGGTGAGGGCCACCGAGCAG	33622213
Qy	808	GT	CGACGGAATTTGGCGGAAAGTTCCAGTTGCCCCGTTAACACCACTCGATGGAATGGG	867
Db	3362212	CT	CGGGAGCGCTGGCCGAGCTGACCGGCATCCCGGTGTGCACAGCGTGAATGGCCCGGCG	3362153
Qy	868	GCT	TTTGACGAAACCAATCCCTTTCGGTGGGTATGTGGGTATGCAATGCGCACCGCTAT	927
Db	3362152	GC	TTTTCCCGACAGCCACCGGCAAACTCGCGCATGCCCCGACATGCAAGGACGGGTGGCC	3362093
Qy	928	GCCA	CTTTGCGGTCAAGCAATGTGATTTGTATGTCAGTGGGGGCCCGTTTGACGAC	987
Db	3362092	GCC	TGGGGGGGCGCTGACAGCGACGAGCACTGTGTATCGCGCTGGTACCCGCTTCGAGAC	3362033
Qy	988	CGG	GTATCGTGGCAACTAGACGAATTTGTCAGCCGCGCAAGATATTCATTTGACATC	1047
Db	3362032	CGG	GTGACCGGCAAGCTGCACTTCGTTGCGCCCGGAAGCCAAAGTCAATCCAGCCGACATC	3361973
Qy	1048	GACC	CGGAGGTGGGAAAAAAGAGGCTCCGATGTGCCATTTGTGGGGATGTAGCG	1107
Db	3361972	GACC	CGGCGAGATCGGCAAGCCGCCACCGACGNCCTCATGTGGTGAAGTCAAG	3361913
Qy	1108	CAT	GTTTTGAACAGCTTTTGGACGGGGCCCGGGAATTTGATTACCCACCCATCCCAT	1167
Db	3361912	GCC	GTCAACACCGAATCGATGCGATGTGGCCACCAACCATTTCCCGGACCATGAG	3361853
Qy	1168	AC	CAACCCGAGCATGTTAATTCGATTTGATTTGGCGACCCGATTAACCCCTCCAGATG	1227
Db	3361852	AT	GGCCCACTGTGGCATTACTCTGAAGGTGTGGCAAGAATCTACGCTGAGCTAATGGG	3361793
Qy	1228	CCCC	CTATGAGATACTATTGCCCCCGGAGAGTGTAGTACGAAATTTGTGTGCCAGGCC	1287
Db	3361792	CCG	CAGACGACCGGAGCTGAGCCCGGAAATCGAATCGAAAAAGCTCGGAGATGGCC	3361733
Qy	1288	---	CCCCATGCTCTATCACCAACCGAATGTGGACAAACCAAAATGGGGCGCCAGTTT	1344
Db	3361732	GGG	CCGGAGCGCGATTTGTGTGCGGGCGCTCGGCGACACAGATGTGGGCCGCCAGTTCC	3361673
Qy	1345	TTGA	---ACAATGGCCCCCGCGCATGTGAATTTTCAAGTGTGGCTTGGTACAGTGGCTTT	1401
Db	3361672	AT	CAGATTCGAAAAAGCCGCGCAGCTGGGTGAATCCGGCGGTCTGGGACCATGGGCTTT	3361613
Qy	1402	GTT	TACTGCGCGCATATGGGAGCCAAATGGGAAT - GGGGAGCAGAGCGTTCATTGGCATC	1460
Db	3361612	GCCA	TCCCGGGCCATATGGGCGCCAAAGTGCCTCCCCCGGACCGAGATCTGGGCGATC	3361553
Qy	1461	AGT	GAGATGCGCAGCTTCCAAATGATTTTAGAGACTGGGAACTTACCCCACTAGCAGC	1520
Db	3361552	GAC	GGCGGACGATGTGCTTCAGATGACCAACAGAGAGCTGGGCACCTGGGGGTGCGAGGCG	3361493
Qy	1521	AT	CAAGGTTAAACTATTATTCTCAATACGCTTGGCAGGGGATGGTGTCTCATGTGGCA	1580
Db	3361492	AT	ACCGGTCAAGGTGGCGCTATCAACAACCGCAACTGGGCATGTGTGCGCAATGTGCAG	3361433
Qy	1581	CA	ACTTCTCTACGAAGAAGTATTCTGTCTTAACATG-----TCCAGGGCGATG	1631
Db	3361432	AG	CGTTCTATTCGCGAGCGGTACTTCGAGACCGAGCTGGCACTCATTTGGCACCGCATC	3361373
Qy	1632	CC	AGCATTTAATCTCTCTGTGAAGCCTATGGCATCAAGGGTATTACTGTGCGCAAGCGG	1691
Db	3361372	CCG	CGCTTCGTAAGAACTGACGAGGCTTTGGGGTGTGTGGGTGTGCGGTGCGAGCGGGAA	3361313
Qy	1692	GAA	GATTTGGCCCCCGGCATGCGCCAAATGTGTAGC---CCACAAATGTCTGTGTGATG	1748
Db	3361312	GAG	AGCGTGTGTGAGTGCATCAACCAAGGCGGGGGGATCAACGACTGCGCGGTGTGATC	3361253
Qy	1749	GAT	GTGTGTGTCAAAAAAGATGAATCTGTATACCTATGATTTGCCCCCGCATAGTAT	1808
Db	3361252	GAC	TTCATCTGTGTGTCGCGAGCGCAATGTGTGGCCGAATGTGTGGCGCGGACCAAGCAAT	3361193
Qy	1809	GCCC	AAATGC 1818	

	RESULT 6	
US-09-103-840A-2/c	/ Sequence 2, Application US/09103840A	
Patent No. 6294328		
GENERAL INFORMATION:		
APPLICANT: FLEISCHMAN, Robert D.		
APPLICANT: WHITE, Owen R.		
APPLICANT: FRASER, Claire M.		
APPLICANT: VENTER, John C.		
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM TUBERCULOSIS		
FILE REFERENCE: 24366-20007.00		
CURRENT APPLICATION NUMBER: US/09/103,840A		
CURRENT FILING DATE: 1998-06-24		
NUMBER OF SEQ ID NOS: 2		
SOFTWARE: PatentIn Ver. 2.1		
SEQ ID NO 2		
LENGTH: 4403765		
TYPE: DNA		
ORGANISM: Mycobacterium tuberculosis		
FEATURE:		
OTHER INFORMATION: CDC 1551		
OTHER INFORMATION: "n" bases at various positions throughout the sequence		
OTHER INFORMATION: represent a, t, c or g		
US-09-103-840A-2		
Query Match	18.7%; Score 365.4; DB 3;	Length 4403765;
Best Local Similarity	53.8%; Pred. No. 2,1e-108;	
Matches 942; Conservative 1;	Mismatches 773;	Indels 34; Gaps 8;
Db		
Oy	88 CCTATATCCCAAAACAAATTTCCTGAAAATGTCTTCTAGCCAACGGCAAACCGGGCCTTAT	147
Db	3357249 CCAAGACACC GGCGGCTCGCTCGAAACATGTTGCATCTGACAGACGCTTACCGGTGCACAG	3357190
Oy	148 ATCTCATGTGATGATGCCCTGAAGAAGCGCATGGGCTCAACAACATTTTTGGCTATCCCGGCGG	207
Db	3357189 GCGETCATCCGCTGCTGAGAGAACTCGCGCTGCACCTCATTTTCGGGATTTCCGGCGGT	3357130
Oy	208 GCATTTTGGCCCATCTATGTATGAACTGTACCGCTTTTGAGACGGCGGGGAATTTGACAT	267
Db	3357129 GCGGTGCTGCGCGGTGATGACCCGCTGTTCGACTCGAAAAAGCTGCSC-----CAC	3357079
Oy	268 ATTTTGGTGGCCCATGAAACAAGAGCTTCCCATGCGCGGATGGGTATGCGCAGAGCCACA	327
Db	3357078 GTGTGATGCGCCACGAAACAGGGGCGGGCATGCGCCACGGGCTATCGCACCTCAC	3357019
Oy	328 GGTAAATGGAGATTGTTTTCGTGATCTGGAACAGGGGCGCACTAACTTGTGACCGGC	387
Db	3357018 GCGCGGATGGCGGTGTGCATGCAACGTCCGGATCCGGAGACAAACCTGTGATACCCCG	3356959
Oy	388 ATTGCCAATGCCCATTTTGAAGTCTGGTCCCATGTTGTTGATTAATCTGAGAGGTGGCCGT	447
Db	3356958 CTGGCCGACGGCGAGATGAGACTCGATCCCGGTGTGCTATACCGGTCAGGTGCGGCGG	3356899
Oy	448 GCCATGATGGTATGCGATGCTTTCAGAGAAATTGACATTTTGGCATCACCTAACGCATC	507
Db	3356898 GGGCTGATCGGACCGACGCGCTTCCAGAGAGCCGACATCTCGGGCATCAAGATGCCATC	3356839
Oy	508 GTTAAGCACTCTATGTGATACGTATGTCGCGGATATGAGCTGCAATGTTACTGAGGCT	567
Db	3356838 ACCAAGACAACTTTTGTGTCGCTCGGTGACGACATTCGCGGGGTCTGGCGGAGGCGC	3356779
Oy	568 TTCCATCTTGTGATGACACCGGTGTCCTCCGGGCGGTTTGTATTCATATATCCCAAGAGATG	627
Db	3356778 TTCACATCTGGGCTCCGCGCGCGGCGGGGTGTGTGATGACATCCCCAAGAGAGTG	3356719
Oy	628 GGCTTAGAAGATGTGATCATATCCCCTCGACCCCGGTATGATTAATCAACCGGGTTAT	687
Db	3356718 CT----GCAAGGCGCAGTGCACGTTTCACTGCGCGCGCGGATG--GAGCTGCGCGGCTAC	3356655

```
QY      688  CCCCCCGGTTAAAGGTAATCCCGACAAATTAAATGCGCGATTGSCAATTGTTGAGCAG  747
Db      3356664  AAGCCCAACACCAAAACCGGACAGCGCGAGGTCCCGGAGGCGCCCAAGCTGATCCGGCC  3356605
QY      748  GCCAGAAATCCCTGTGCTACAGTAGAGGGGAGGGGAGTGGCGCAATGCCATGCCATCCGAG  807
Db      3356604  GCGCCGACGCCGCTGTCTGTATGTCCGCGCGCGCTCATCTCGGTGAGGCCACCGACGAG  3356545
QY      808  GTGCAGGAATTTTGCAGAAAGTTCCAGTTGCCGATTAACACCACTGATGGAATTTGGG  867
Db      3356544  CTCGCCGAGCTGCGCGAGCTGACCGGCATCCCGGTGTGCACACAGCTGATGGCCCCGGCG  3356485
QY      868  GCTTTTGAAGAAACCATCCCTTTCCGTGGGTATGTTGGGTATGGATGGCCACCGCTAT  927
Db      3356484  GCGTTCCCGACAGCCACCGGAAAACTCAGCAATCCCGGATGCAACGCGATGCGC  3356425
QY      928  GCCAATTTCGCGTACGCAATGATGATTTGTTGATTGAGTAGTGGGGCCGTTTGCAGCAG  987
Db      3356424  GCCGTGGCGCGCTGAGAGCGACGACCTGTGATCGCGTGGTATCCCGCTTGCAGCAG  3356365
QY      988  CCGGTAACTGCGAAACTAGACGAATTTGCTAGCCCGCCAAAAGTAATTCAATTGACATC  1047
Db      3356364  CCGGTGACCGCGCACTGACCTGCTTCGCGCGGAAAGTCATCAGCGCGACATC  3356305
QY      1048  GACCCGCGAGAGTGGGAAAAACAGGGCTCCGATGTGCCATTTGTGGGGATGTACGC  1107
Db      3356304  GACCCGCGCGAGATGCGCAAGAACCCGACCGGACGTGCCATGTCGTGATGCTCAAG  3356245
QY      1108  CATGTTTGAACACGCTTTTGCAGCGCGCCCGGAAATTGATTAACCCCAACCATCCCAT  1167
Db      3356244  GCGGTATCACCCGAACTGATCGCGCTGCGCCACCAACATCCCGGACCATCGAG  3356185
QY      1168  ACCACCCGAGCATGTTAAATGCGATTGATCATTTGGCGGACCGATTACCCCTCCAGTG  1227
Db      3356184  ATGCGCGACTGTGTGGCACTACTTGAAAGGTGTGCGCAAGCTTACCTGATGTAGG  3356125
QY      1228  CCCCCTATGAGGATCTAATTTGCCCCCGAGAGGTAGTACAGAAATGTGTGCGCAGGCC  1287
Db      3356124  CCGCAGAGCGAGCGGACGCTGAGCCCGGAATACGTGATCGAAAGCTCGGGAAGATGGCC  3356065
QY      1288  ---CCCGATGCTACTTACACCAACCGATGTGGGACAAACCAAATGTGGCGGCCAGTTT  1344
Db      3356064  GGGCGGAGCGCGTGTGTGTGCGCGCGCTCGCCGACCAAGATGTGGCGGCCAGTTC  3356005
QY      1345  TTGA---ACAAATGSCCCCGCGCATGATGATTTCCAGTGTGCTTGGGTAGCATGGGCTTT  1401
Db      3356004  ATCAATATGAAAAAACCCCGGACGTGTGTAATCCCGCGGTGTGGGACCATGGGGTTT  3355945
QY      1402  GGTTAATCTGCGCGCATGAGGACCAAGTGGAGT-GGGGGACAGACCGGTCAATTGATC  1460
Db      3355944  GCCATCCCGCGCGCATGAGGCGCCAAAGATCCCGCTCCCGGACACGAGGTCTGGGCATC  3355885
QY      1461  AGTGAGATGCGAGCTTCCAAATGAACTTTCAGGAACTGGGAACTTATGCCAATGAGAC  1520
Db      3355884  GAGCGCGACGTTGCTTCCAGATGACCAACAGAGCTGGCCCACTGCGCGGTGTGAAGGC  3355825
QY      1521  ATCCAGTTAAACTATTAATTTCTCAATAACGTTTGGCAGAGGAGTGTGCTGCAATGCA  1580
Db      3355824  ATACCGGTCAAGTGGCGCTGATCAACAACGGAACCTGGGCAATGTGTGCGCAGTGGAG  3355765
QY      1581  CAAATCTTCTGGAAGACGTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  1631
Db      3355764  AGCTGTTCTATGCGGACGCTACTCGCAACCGACCTGSCCACTCAATTGCAACCGATC  3355705
QY      1632  CCAGACATTAATCTCTCTGTGAAAGCTTATGAGCATCAAGGGTATTAATGTCGCAACGG  1691
Db      3355704  CCGGACTTGTGAAACCTGCGCGAGGCTTGGGGGTGTGTGGGGTGTGGAGCGGGA  3355645
QY      1692  GAAAGTTTGGCCCGCGCATGCGGAAATGCTAG---CGCAATGTGCTGTGTGTATG  1748
Db      3355644  GAGGAGTGTGTGAGCATCAACCGAGCGCGGGGATCAACGATGCGCCGCTGTATC  3355585
QY      1749  GATGTGTGTGCAAAAAAGATGAATACTGTATACCTATGATTTGCCCGCGCATGAGTAAT  1808
```

```
Db      3355584  GACTTCATGCTCGGTGCGGACGCGCAAGTGTGCGCATGAGTGGCGCGGCGCACCAAGAT  3355525
QY      1809  GCCCAATGC  1818
Db      3355524  GACGAGATCC  3355515

RESULT 7
US-09-540-236-646
; Sequence 646, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT?
; FILE REFERENCE: 2709, 2005-001
; CURRENT APPLICATION NUMBER: US/09/540, 236
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 646
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-646

Query Match      18.0%; Score 353; DB 4; Length 1632;
Best Local Similarity 54.8%; Pred. No. 1.2e-106;
Matches 824; Conservative 0; Mismatches 650; Indels 31; Gaps 5;

QY      137  CCGGGGGTTATTCGATGAGTATGACCTGAAGCCATGAGGGGTCACAACTTTTGGCT  196
Db      122  CAGCGCTGATGATGCTTTTCAAGCATGATGATGATGATGATGATGATGATGATGATGAT  181
QY      197  ATCCCGCGGGGCAATTTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGAT  256
Db      182  ATCAGGCGGTGTGGTATTAATATTAATGATGATGATGATGATGATGATGATGATGATGAT  232
QY      257  AAATTGAGCATTTTGTGTGCGGCATGAACAGAGACTTCCATGCGCGGATGGTATG  316
Db      233  AAATTGAACATTTTGTGTGCTGATGAACAGCCAGGCGATATGCGCGATGATGATGATGAT  292
QY      317  CCAGAGCCACAGGTAAGTGGAGTTGTTGCGTATCTGATGACCAAGGGCGCATGACT  376
Db      293  GTGTGAACAGGACAGAGAGGGGTGTTTGGCGACTTCAGTCCAGTGCAGTCAACAA  352
QY      377  TGGTACCGGATTTGCCAATGCCATTTGACTCGGTGCCATGTGTGTGATTACTGAG  436
Db      353  CTGTAACAGCATTTGCACTGCTTATATGATTAATCCCAATGTTGTGTGGCAGTTC  412
QY      437  AGGTGGCGCGTGCATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT  496
Db      413  AGGTACCATCTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  472
QY      497  CTTTACCATGTTTGAACATCTCTATGTGTGATGATGATGATGATGATGATGATGATGAT  556
Db      473  CACGCCCATCTGTTAGCAACAGTTTCAAGTTGCGCATCCGGAAGACATTCACAAATCA  532
QY      557  TTTCTAGGCTTTCCATCTTGTGAGACCGGTGTGCTCCCGCGCGGTGTTGATGATGAT  616
Db      533  TCAAAAAGCATTTTATTTATGAGCTCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  592
QY      617  CCAAGATGTGGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  676
Db      593  CAAAGATATGACCAACCAAGATTAATTTGCTTATATCAATGCCAAGACATCAATA  652
QY      677  TACCGGTTATGCGCCCAAGGTTAAGATATCCCGAACAATTAATGCGCATGATGAT  736
Db      653  TCAGATCATATCAAGCTTCAATTAAGATGATGATGATGATGATGATGATGATGATGAT  712
QY      737  TGTGAGAGAGCGCAAAATCCCTTGTCTATGATGAGGGGAGGGGATGCGCCGCAATG  796
Db      713  CGTACTATCAAGCAACGCCCTGTATCTGATTAATTCAGTGTGTGTGTGTGTGTGTGTGT  772
```

797 CCCATGCCAGTGCAGGAAATTTGGGAAAGGTTCCAGTTGCCGATACACCCCTGA 856
773 CCACTAAGAGCTGCTGATCAATTAAGCCATAGCTCAATCCTGTTACTTAATACCTTA 832
857 TGGGAATTTGGGCTTTTGAAGAAAACATCCCTTTGGGGTATGTTGGATATGATG 916
833 TGGAGCTGGGCTTATTCAGAGGCTCAGGCGACATTTGTTGGATCTGTATGATG 892
917 GCCACCCGTATGCCAATTTGGCGCTACGCGAATGATTTGATTTGATGAGTGGGGCCC 976
893 GCACTATTAAGGCAACATGACATGATCTGCTGATGATTTTGGCAGTGGGTGTC 952
977 GTTTCAGCAGCCGGGTAATGGAACATAGCAATTTGCTAGCCCGCAAGTAATTC 1036
953 GTTTCATGACCGTGTGACCAATATGTCATAAATTTCTGCCCAATGACATCATTC 1012
1037 ACATTTGATGACCCCGGCGAGGTGGGAAAAACAGGCTCCCGATGTCCTATGTCG 1096
1013 ATATTGATGATGATCCAGCCAGTATCTCAAGAGATTAATGACACATTCGATTTG 1072
1097 GGAATGATGCGCATGTTTATGACAGCTTTTGAAGCGCGCCGGAATGATTAACCA 1156
1073 GGAATGATGATGATGATTAATGATGATGATGATGATGATGATGATGATGATG 1132
1157 CCATTCCTCATACACCCAGGAGGATTAATGCAATGATGATGATGATGATGATG 1207
1133 TCGATGACATGCTCTCATGATTTGGGCAAAATCAATGATGATGATGATGATG 1192
1208 -----CCGATTACCCCTCCAGGTGCCCACTATGAGATATCTATTTGCCCGAGAG 1261
1193 GTCTTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252
1262 TAGTACAGCAATTTGTTGCGCAGC---CCCGATGCCCTACTACACACCGATGTCGAC 1318
1253 TGGTGAAGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1312
1319 AAGCAAAATGTTGGGCGGCGCCAGTTT---TGAACATGAGCGCGCGATGATTTCCA 1375
1313 AGCATCATGTTGTTGCTCATGTTATTAATGATGATGATGATGATGATGATGATG 1372
1376 GTCTGCTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1435
1373 CAGGTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1432
1436 TGGGGGAGCAG---CGTCACTTTGCTCATGATGATGATGATGATGATGATGATG 1494
1433 ACCCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1492
1495 AACTGGGAACCTTAGCCAGTACGACATCCAGTTAAATCTATTTCTCATATCCGTT 1554
1493 AGCTGTCAATTTGCTCATGATGATGATGATGATGATGATGATGATGATGATG 1552
1555 GCGAGGGATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1614
1553 AGCTTGGATGTTAAGCAGTGGAGGATGATGATGATGATGATGATGATGATG 1612
1615 ACATG 1619
1613 ATATG 1617

RESULT 8
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and uses thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186F3
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 16.8%; Score 328.4; DB 4; Length 1830121;
Best Local Similarity 53.1%; Pred. No. 2.3e-96;
Matches 771; Conservative 0; Mismatches 667; Indels 13; Gaps 3;
173 ATGGGGTGAACACATTTTGGCTATCCCGGGGCGCATTTTGGCCATCATGATGATG 232
1652874 AAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652816
233 TGTACCGCTTGAACCGGCGGGGAAATGATGATGATGATGATGATGATGATGATG 292
1652815 -----ANTCATCGCTGGGTGATGATGATGATGATGATGATGATGATGATGATG 1652764
293 CTTCATCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 352
1652763 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652704
353 CATCTGACCAAGGGCGATCTATGATGATGATGATGATGATGATGATGATGATGATG 412
1652703 CTTCGGGCGCTGTCACAAATGCTATGATGATGATGATGATGATGATGATGATG 1652644
413 TCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 472
1652643 TCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652584
473 AGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532
1652583 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652524
533 GTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 592
1652523 AGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652464
593 CCGGGCGGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 652
1652463 CAGGTCTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652404
653 CCTGACCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 712
1652403 ATGATATCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652344

Mon Jul 26 12:13:49 2004

us-09-893-033-6.rn1

Page 9

713 GACAAATTAAATGCGGATTCGATTTGAGAGAGCCAGAAATCCCTTGCTACGATG 772
Db 1652343 GTCAATTAATAAAGCGTTAAAGACATTTAGTGGCGAAGAAACAAATACCTTTTTCG 1652284
QY 773 GGGAGGAGGCGATCCCGCCCAATGCCCAGGTGCAAGAAATTTGCGAAAAGTTCC 832
Db 1652283 GGGGTGTGCAATCACTGTGATGTAGTGAAGCAATTAATCAGTTTGCAAGCTTAA 1652224
QY 833 AGTTCGCGGTGACAAACACCTGATGGGAATTTGGGGCTTTTACGAAAACATCCCTTT 892
Db 1652223 ATTTACCGGTGATCTCTCATTAATGGGATTTGGGCTTATCAAGTACGATTAACAT 1652164
QY 893 CGGTGGTATGTTGGGTATGATGACGACCGGTATGCCATTTGGCGTACGCAATGTG 952
Db 1652163 TCTTAGGTATGTTGGGTATGATGATCTTAAAGCAATACGCAATGACCAAGTGT 1652104
QY 953 ATTGCTGATGACGTGGGGGCGCGCTTTCGACGACCGGGTAACTGGCAACTAGACAT 1012
Db 1652103 ATCTTATTTTGGGGATGGCGCTTCTTTGATGATGATGACGACAAATATTAGAAAAT 1652044
QY 1013 TTGCTAGCGCGCCCAAGTAATTCATTCATGACATGACCGCGCGAGGTGGAAAAACA 1072
Db 1652043 ATTGCCCCAATGCAAAAGTATTCATTAATGATGATCAACTCAATTTCTAAAAATG 1651984
QY 1073 GGGCTCCCGATGTCGCAATTTGGGGGATGTACGCGCATGTTTAAACAGCTTTTGACG 1132
Db 1651983 TGCCAGTATGCGATTCATTAATGCGAAATGCGAAAATGTATGGAAGAAATTTTGGGTT 1651924
QY 1133 GGGCGCGGGAATGGATTAACCCCAACCATCCCATACCAAGCATGTTAAATGCA 1192
Db 1651923 TATTGAATGAAGAGGATTAATTAATTCAAACAGTCTTGAAGTGGGCGCAAAATCA 1651864
QY 1193 TTGATCATTTGGCGGACCATTAACCCCTCCAGGTGCGCCCATATGAGGATCTATGCCC 1252
Db 1651863 ACCATGGAAGAAACAAAATGTTTGAATTTGACCGCATCTTGCGGTATTAACCGC 1651804
QY 1253 CCCAGGAGTATGATACGAAATTTGTCGCGACGCGCGCGATGCTTACACACCGATG 1312
Db 1651803 AACAGTGTGGAAGCGGTATATGCGCTCAGAAAGACAGCTTATGTGCTTGATG 1651744
QY 1313 TGGGACAAACCAATGTGGCGGCGC---CGAGTTTGAACAATGGCCCCCGCGATGGA 1369
Db 1651743 TAGGTGACACCAAAATGTTTGTGCTTACATTAATTCATTTGATGAACCGGTATGGA 1651684
QY 1370 TTTCCAGTGTGCTTGGGTACGATGGGCTTTGTTTACCTGCCCATGGAAGCAAG 1429
Db 1651683 TTAATTTGTTGGTACGACCAATGTTTGTGCTTACATTAATTCATTTGATGAACCGGTATGGA 1651624
QY 1430 TGGGAG-TGGGGAGAGAGCGGTATTTGATCACTGATGAGATGCCAGCTTCAATGATC 1488
Db 1651623 TAGCTCATCTGGAAGAACTGTGTGCTGCTTACTGCGAGTGAATTAATCAATGATA 1651564
QY 1489 TTCAAGAACTGGGAACCTTAGCCGATGACATCCAGTTTAAACTATTAATTTCTATA 1548
Db 1651563 TTCAAGAACTTTCAACGCAACACAAATGTCATTCCTGTTGTCAATTTTGTGATA 1651504
QY 1549 ACCGTTGGCAGGGGATGTTGCTGATGATGCAACAACTTTTACGAAGAACTTATTTG 1608
Db 1651503 ATCAATTTCTTGGGAATGTTAAACAAATGCAAGATTTGATTTATTTCTGCGCATTTGCG 1651444
QY 1609 CTTCTAACATG 1619
Db 1651443 AAACCTAATG 1651433

RESULT 9
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams

Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643, 990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
Query Match 16.8%; Score 328.4; DB 4; Length 1830121;
Best Local Similarity 53.1%; Pred. No. 2.3e-96;
Matches 771; Conservative 0; Mismatches 667; Indels 13; Gaps 3;
QY 173 ATGAGGCTCAAAACATTTTGGCTATCCCGGCGGGCAATTTGCCCATCTATGATGAAC 232
Db 1652874 AAGCGTAGAGATGATTTGTTATCCCGAGGCGCGGTATTAATTTATGATGCA- 1652816
QY 233 TGTACCGTTTGAAGCGCGGGGGAATAGACATATTTTGGGCGCCATGAACAAGAG 292
Db 1652815 -----ANTCATAGCTGGGTGTGATTTGAACATATTTTATGTTGCCCAACAAGCCG 1652764
QY 293 CTTCCCATGCGCGGATGGGTATGCGACAGCCACAGTAAAGTGGAGTTTTCGGTA 352
Db 1652763 CGGTGATATGCGCGGATGGTTACGACGTTCAACGGGTAAAGTTGGATGTGCTAGTA 1652704
QY 353 CATCTGACACGAGGGGACTAATCTTGATGACCGGATTTGCCAATGCCATTTGCACTCGG 412
Db 1652703 CTTGCGGCGCTGGTGCABAAATGCTATCACTGGCATTTTAAACGCTTATACCGATTCG 1652644
QY 413 TGCCATGTGTGATTTACTGAGAGAGTGGCGGTGACATGATTTGATGCAATGCTTTCC 472
Db 1652643 TGCTATGTATATCTTTCGAGGTCAAGGTATAGCAATTTATTTGCGAGTATGCTTTCC 1652584
QY 473 AGAAATTAACATTTTGGCATCACCTTAACGATGTTTAAGACATCTTATGTGATGTA 532
Db 1652583 AAGAAATGATATGCTGGGATTTTCTGCTCCGTTGTGAAGATGCTTTATTTGCAAA 1652524
QY 533 GTGCGCGGATATGCTGCTATTTACTGAGAGCTTTCCATCTTCTGACACCGGTGCTC 592

Db 1652523 AGCGGAGATATTCATCTCTTGAAGAAAGCCTTTATATTCATCTACGGGTGTC 1652464

QY 593 CCGGCGCGGTTTATGATGATATTCACAGATGCGGCTTGAAGATGAGTACCTTC 652

Db 1652463 CAGGCTCTGTGTGTGATATTCGAAAGACCTGTAAATCTTAATTTTAAATCTTC 1652404

QY 653 CCGTCGACCGCGGATGATATCTACCGGGTATGCGCCCGAGGTTAAAGTATCCCC 712

Db 1652403 ATGATATCCCTGAATATGTGAGTACCTTCTTATATCCAGAGTAAAGACAAAG 1652344

QY 713 GACAAATTAATGCGGATGATATGATGAGAGCGCCGAGAAATCCCTGCTACGTAG 772

Db 1652343 GTCAATTAATAAAGCGCTTAAAGACACTTTAGGCGGAGAAACCAATCTTTTGTG 1652284

QY 773 GGGGAGGGGCGATGCGCGCAATGCCATGCCGAGGCGAGGATTTGGCGAAAGTTC 832

Db 1652283 GTGGGAGCATATCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1652224

QY 833 AGTTGCGGTAACACACCCCTGATGAGGATTTGGGCGCTTTGACGAAACCATCCCTTT 892

Db 1652223 ATTTACCGGATCTTCCTCATTAATGAGATGCGGTGCTTATCCAGATGAGATTAACAT 1652164

QY 893 CCGTGGGATATGTTGGATGATGATGATGATGATGATGATGATGATGATGATG 952

Db 1652163 TCTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652104

QY 953 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1012

Db 1652103 ATCTTATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1652044

QY 1013 TTGCTAGCGCGCGCAAGTATTAATCAATGATGATGATGATGATGATGATGATG 1072

Db 1652043 ATTGCCCAATGACAAAGTATGATGATGATGATGATGATGATGATGATGATG 1651984

QY 1073 GGGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132

Db 1651983 TGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651924

QY 1133 GGGGCGCGGAAATGATTAACCCACCAATCCCATACCAACCGGATGATTAATGCA 1192

Db 1651923 TATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651864

QY 1193 TTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252

Db 1651863 ACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651804

QY 1253 CCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1312

Db 1651803 AACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651744

QY 1313 TGGGACACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1369

Db 1651743 TAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651684

QY 1370 TTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429

Db 1651683 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651624

QY 1430 TGGGAG-TGGGAGACGCGGATGATGATGATGATGATGATGATGATGATGATG 1488

Db 1651623 TAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651564

QY 1489 TTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1548

Db 1651563 TTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651504

QY 1549 ACAGTGGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1608

Db 1651503 ATCAATTTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651444

QY 1609 CTCTAATG 1619

Db 1651443 AACCTATATG 1651433

RESULT 10
US-09-489-039A-840
; Sequence 840, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 840
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-840

Query Match 16.5%; Score 323.8; DB 4; Length 1689;
Best Local Similarity 53.4%; Pred. No. 6.5e-97;
Matches 885; Conservative 0; Mismatches 702; Indels 70; Gaps 7;

QY 172 CATGGGATCAACACATTTTGGCTATCCGCGGCGCAATTTGCCATCTATGATGAA 231

Db 82 CAGGAGATGACACAGTTTGGCTATCCGGTGGCGCAATTAAGCGGTTACGATGCA 141

QY 232 CTGTACCGCTTTGAAGCGCGCGGGAATGATGATATTTGGTGGCGCATGACAGAA 291

Db 142 CTGTATGAC-----GGCGCGTGAAGACCTCTGTGCGGACATGACAGAAAGG 189

QY 292 GCTTCCATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351

Db 190 GAGCGATGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 249

QY 352 ACATCTGAACAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 411

Db 250 AGCTCGGCGCGCGGCGACATGATGATGATGATGATGATGATGATGATGATGATG 309

QY 412 GTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 471

Db 310 ATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 369

QY 472 CAGAAATGACATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 531

Db 370 CAGAAATGACATTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 429

QY 532 AGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591

Db 430 TCTCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 489

QY 592 CCGGCGCGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 651

Db 490 CAGGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 549

QY 652 CCCCTGACCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711

Db 550 CCCCATTTCTTACCGTGGCGGATGATGATGATGATGATGATGATGATGATGATG 595

QY 712 GCACAAATTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 771

Db 596 ----AAAGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 651

QY 772 GGGGAGGGGCGATGCGCGCATGATGATGATGATGATGATGATGATGATGATGATG 831

Db 652 GCGGCGGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711

QY 832 CAGTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 891

Db 712 AGGATCCCGGCAACTGCACTCCGTAAGGAGGCTGGGATGCTGATTAATCCCTAT 771
 QY 892 TCGGTGGATATGTGGGATGATGATGAGCCACCCCTATGCAACTTTGCGCTGACGATAT 951
 Db 772 TATCTGGCATGTGGGATGACAGGCACTAAGGCGGCCAAGCTGCGGTATAGAGAGTGC 831
 QY 952 GATTTCTTATGATGCACTGGGGGCGCTTTGACAGACGCGGTAATCTGGCAACTAAGCAA 1011
 Db 832 GATCTGCTATGCGCGCTGCGCGCGCTTTTGAAGATCGGTAACCGGTAACTCAATACC 891
 QY 1012 TTTCGTAAGCGCGCCCAATTAATTCATTTGATGATGACCGCGGAGAGTGGGAAAAAC 1071
 Db 892 TTGCAACCGCATGCGCAAAATGATCCATGACATGATCCCGGAGCTGAACAAGCTG 951
 QY 1072 AAGGCTCCCGATGTGCCCATTTGTGGGAGTGTACCGCATGTTTAAACAGCTTTTGAG 1131
 Db 952 CCGTACAGCGCATATCGGTCTGACAGGCGATCTGAATGCTGCTGCTGCGCGCTGACAG 1011
 QY 1132 CCGGCGCGGGGAATGGAATTACCCCAATCCCAATCCCAAGGCAATGTTAATAGC 1191
 Db 1012 -----CGTTAGCTATGATGGCGCGAGCGG 1041
 QY 1192 ATGTATCATTTGCGGACCGATTAACCC--CTCAAGTGGCCCACTATAGGATCTAT 1248
 Db 1042 AGTGGCGCATGTGCTGCGAGCATGCTGGCGTTAGATCCCGGTAGGCAATCTAC 1101
 QY 1248 GCGCCCGGAGAGTGTACACGAATTTGTGCGCAGGCGCGCGATGCTCTTACACCAAC 1308
 Db 1102 GCGCGCTGCTCTGAAGAGCTTTCCAGCCGCAAGCGCGAGAGGTGTGACCAAC 1161
 QY 1309 GATGTGGGACACACCAAAATGTGGGCGGCCAGTTTGAACATGCGCCCGCGCA--- 1365
 Db 1162 GACGTGGGCAACACCAAGATGTGTGCTGACGATGATGATCTAATCCCGCGGAAAC 1221
 QY 1366 TGGATTTCCAGTGTGCTTGGGATGATGAGGCTTTGGTTTACCTGCGCGCATGGAGCC 1425
 Db 1222 TTATACCTCCAGCGGCTTAAGGACCAATGGGCTTTGGGCTGCGCGCGCTGGCGCC 1281
 QY 1426 AAGTGGGAGTGGGGGAGCA-GGGTCATTTGATCAGTGAAGGAGCCAGCTTCCAAATG 1484
 Db 1282 CAGGTGGCGCGCCCTTAACGATACGATGATCTGATCTGCGGTGAGCGCTCTTCATGATG 1341
 QY 1485 AATCTTCAGAACTGGGAACTCTTAAGCCAGTACGACATCCAGTGAATTAATTTCTC 1544
 Db 1342 AATGTACAGAGACTGGGACCGTTAAAGCAGGATTAACGCTGAAGTGTGTATCTC 1401
 QY 1545 AATAACGTTGGCAGGAGTGTGTGCTCAGTGGCAAACTTTCTAGAGAAAGCTTAT 1604
 Db 1402 GACAAACAGCGTTAAGGATGTGTGACATGACAGCGCTGTTTTCAGAGAAAGTAT 1461
 QY 1605 TCTGCTTCAATGATCCAGGAGCATGCGACATTAATCTCTCTGGAAGCTTATGCG 1664
 Db 1462 ACGGAACCACTGTAC--TGATTAACCTGATTTCCAGCGCTGGCAGCGCTTGGCT 1518
 QY 1665 ATCAAGGATTAATCTGTGCGCAGCGGAAAGTTGGCCCGCGATGCGGAATGTCTA 1724
 Db 1519 ATTCCTGGCAACACATCAACCGTAAAGCCAGTGAAGCGGCACTGACCAATGCTT 1578
 QY 1725 GCCCAATGCTGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1784
 Db 1579 TCGAACCGAGGCGCATCTGCTTCAATGATGATGATGATGATGATGATGATGATGATG 1638
 QY 1785 ATGATTTGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
 Db 1639 TTGTGCGCGCGCGCGCGCATGATCTGAAATGCTGG 1675

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-033A
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 3942
 ; LENGTH: 1731
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-3942

Query Match 16.4%; Score 321.6; DB 4; Length 1731;
 Best Local Similarity 53.7%; Pred. No. 3.6e-96;
 Matches 808; Conservative 0; Mismatches 674; Indels 22; Gaps 6;

QY 128 AAGGCAACCGGCGCTTAATCTGATGATGATGATGATGATGATGATGATGATGATGATG 187
 Db 11 AACTTTATCTGTGTGAATGCTGCTGCGCTTGGGACGAAAGCGTTGAATG 70
 QY 188 TTTTGGCTATCCCGCGGCGCAATTTTGGCCATCTATGATGATGATGATGATGATGATG 247
 Db 71 TTTTGGTTATCCAGCGGCGGAGTATTAATTTATGATGATGATGATGATGATGATGATG 130
 QY 248 CCGGCGGAGAAATGACATATTTTGTGCGCCATGACAGAGCTTCCCATGCGCGG 307
 Db 131 -----AATCTAATCATTAACCTGCTGCGCATGACAGGCTGCTGATATGACAG 181
 QY 308 ATGGGATGACAGAGCCAGTAAGTGGAGTTGTTTGGGATCATCTGACAGAGG 367
 Db 182 ATGCTCTGCGCGCTGACAGTAAAGTGGCTGATGATGATGATGATGATGATGATGATG 241
 QY 368 GCACTAATCTGTGACCGGCAATTTGCCAATCCATTTGATGATGATGATGATGATGATG 427
 Db 242 CAACCAATGACATCAATCCAAATTTGCAACAGCTTATGATGATGATGATGATGATGATG 301
 QY 428 TTAATGAGAGTGGGCGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 487
 Db 302 TGTCTGCGCGCTGATCAATCTGATGATGATGATGATGATGATGATGATGATGATGATG 361
 QY 488 TTGCATCACTTACCGCATGTAAGCACTCTATGATGATGATGATGATGATGATGATGATG 547
 Db 362 TTGTATTTACGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
 QY 548 CTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
 Db 422 CTGCATTTTAAAAAGCAATTTTATGATGATGATGATGATGATGATGATGATGATGATG 481
 QY 608 TCGATTTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 667
 Db 482 TCGATTTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
 QY 668 ACGTATTTCAACCGGCTTATGCGCCCAAGCTTAAAGTATCCCGCAAAATTAATGCGG 727
 Db 542 AAGTAAGATGCGTTCAATCAACCACTTCAAGCGCGCATGCTGCTCAAAATTCGTAAG 601
 QY 728 CATTCATTTTGGAGAGAGGAGCAATCCCTTGTCTATGATGATGATGATGATGATGATG 787
 Db 602 CATTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 661
 QY 788 CCGCAATGCGCATGCGCCAGGTCAGAAATTTGCGGAAAGTTTCAAGTTGCGCGTAACA 847
 Db 662 AAGGAAATGCTTCTGCGTTATTAACGAACTTCCGATTTATCTGATTTACCGTAACTA 721
 QY 848 CCAACCGAATGAGAAATTTGGGCGCTTTGAGGAAACATCCCTTTGGGTGGATGATGATG 907
 Db 722 ACHACTTATGAGGTTAAGGCGATCCCGGCGATGACCAAGTTATTTAGGATTTGG 781
 QY 908 GATATGATGCGCAACCGCTATGCGCATTTGCGCTGACGATGATGATGATGATGATGATG 967
 Db 782 GATATGACGCTATATGAAAGCCAAATATGCAATGACCATGACGATGATGATGATGATG 841
 QY 968 TGGGCTGCTTCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1027

Db 842 TTGGTGGCGGTTTATGATGACCGCGTAAACAAACCCGCAAAATTCGCCCCGAACTCAA 901
Qy 1028 AAGTAATTCACATTCGATGACATCCCGGCGAGGCGGAGAAAACAGGGCTCCCATGTC 1087
Db 902 AAGGATTCATGTCATATGCAACCTCGAACCTTTCAAAACATTAATGCGCATATTC 961
Qy 1088 CCATGTGCGGGGATGTACGCCATGTTTGAACAGCTTTTGACGCGGCGCCGGAATTCG 1147
Db 962 CGATGTTGTGTGATGACGACAGTTCCTTCAAGAAATGTTGGCAGATTAACATTA 1021
Qy 1148 ATTACCCCA---CCATCCCATATACACCCAGGATGTTAAATGCAATGATTCATGGC 1204
Db 1022 ATGTGTCCTAAACCAAAATCTGAAGCATTTGGGCTGTGGTCTCAATCATATGATGCG 1081
Qy 1205 GGACCGATTACCCCT---CCAGGTCCCACTATGAGATATCTATTTGCCCCCGAGAGG 1261
Db 1082 GCAAGTTCATGCTTTGAAATATGAAATTCATCTGATGACATATGAGCCAGAGAG 1141
Qy 1262 TAGTACAGAAATGTTGTCGCAAGGCC---CGATGCTTACTACACCAACGATTCGGAC 1318
Db 1142 TTGTTGAAACTTTATATATTAAGTGAACCAATGCGCATTTATTAATCTGACATAGTTC 1201
Qy 1319 AACACCAATGTTGGCGGCGCC---AGTTTGAACCAATGCGCCCGCGCATGATTCGA 1375
Db 1202 AACATCAATGTTGTTGTCATCTTATTAATGATCAAAAGTCTCGCAATGATCACT 1261
Qy 1376 GTGCTGCTTGGGATGATGAGGCTTTGTTTACCTGCGCATGCGGCAAGTGGAG 1435
Db 1262 CAGGTGCTTGGGACCATGAGGTGTTGTTTACCTTATGACATGAGTGGCAAGCTTCCT 1321
Qy 1436 TGGGGGAGAGAGG-GTCAATTCACATGATGAGATGACAGCTTCCAAATGATTTAGG 1494
Db 1322 TCCAGACCAACAAATGTTGTTGATTAATGAGGTAAAGCTTCAATCAAGTGTATTAAG 1381
Qy 1495 AACTGGGAACCTTACCCGATGACATCAAGTTTAAATCTATTTCTCAATTAACGTT 1554
Db 1382 AACTTCAACCTGTAAACATATGCTTGAATGTAATACTTGCTGTAACAACGTTG 1441
Qy 1555 GCGAGGATGTTGCTGATGAGTGGCAACAACCTTCTACAGAAAGCTTATTCCTCTA 1614
Db 1442 CATTAAGTATGTTGAAACATGACAAATGCAATGAACTATGAAAGTGTCAATCAAGCTTT 1501
Qy 1615 ACAT 1618
Db 1502 ACGT 1505

RESULT 12
US-10-096-571-11
; Sequence 11, Application US/10096571
; Patent No. 6623944
; GENERAL INFORMATION:
; APPLICANT: RIEPING, MECHTHILD
; TITLE OF INVENTION: Process for the preparation of D-pantothenic acid and/or salts th
; FILE REFERENCE: 211499
; CURRENT APPLICATION NUMBER: US/10/096,571
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: DE 10112102.4
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/304,776
; SOFTWARE: Patent in version 3.1
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 11
; LENGTH: 2111
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (8)..(12)
; OTHER INFORMATION:
; NAME/KEY: mutation

; LOCATION: (8)..(8)
; OTHER INFORMATION: Insertion of the base A at position 8
; NAME/KEY: CDS
; LOCATION: (19)..(1665)
; OTHER INFORMATION:
US-10-096-571-11

Query Match 16.3%; Score 318.8; DB 4; Length 2111;
Best Local Similarity 52.8%; Pred. No. 3.5e-95;
Matches 873; Conservative 0; Mismatches 717; Indels 64; Gaps 6;

Qy 172 CATTGGGTAAACACATTTTGGCTATCCCGCGGGGGAATTTTSCCATGATGATGA 231
Db 58 CAGGGGTGAACACCGTTTTCGTTATCCGGGTGGCGCAATTAATCCGTTTACATCA 117
Qy 232 CTGTACCGCTTTGAAGCGCGGGGAAAATGACATATTTTGTGCGCATGAACAGGA 291
Db 118 TTGTATGAC-----GGCGCGTGAGACATTGCTGTCGACATGACAGAGGT 165
Qy 292 GCTTCCCATGCGCGGATGAGTATGCTCAGAGCCACAGTAAAGTGGAGTTTGTTCGT 351
Db 166 GCGGCAATGCGGCTATCGTTATGCTCGGTGTACCGGCAAACTGCGATATGATCGCC 225
Qy 352 ACATCTGACACAGGGGCGACTAATTGTTGACCGGCAATTGCGCAATGCCATTGGAATCG 411
Db 226 ACGTCTGTCGCGGGCGCAACCACTGATTAACCGGCGTTGGGAGCGACTGTATGATCT 285
Qy 412 GTGCCATGTTGTTGATTACTGAGAGGTGGCCGTGCCATGATTTGATGAGATCTTC 471
Db 286 ATCCCTGTTGTTGCAATCACCGGTCAAGTGTCCGACCGTTATGAGACGACGACATTT 345
Qy 472 CAGGAATGACATTTTGGCATCTTACCTTACCGATCGTATTAAGCACTCCATGTGGTACGT 531
Db 346 CAGAGATGATGCTCTGGATTTGCTTACGCTTACCAAGCAAGCTTCTGTTGAG 405
Qy 532 AGTGGCGGATATGAGCTTCGATTTTACTGAGGCTTTTCATCTTTGATGACACCGTCT 591
Db 406 TCGCTGAAGAGTTGCTCGCGCATTTATGCTGAAGATTCGACGTTGCCAGGCTCAGGTGT 465
Qy 592 CCGCGCGGCTTTGATGATATTCACCAAGATGAGGCTTAAAGAAATGATGATCACT 651
Db 466 CCGTGGCTGCTTGTGATATCCAAAGATATTCAGTACGAGGCTGATCGTGA 525
Qy 652 CCCCTGACCCCGGATGACGTTATCTACCGGTTATGCCCCACGTTTAAAGTATCC 711
Db 526 CCGTGGTTACCAACCGTTGAATAAGAGATTTTCCACATGCGG----- 571
Qy 712 CGACAAATTAATGCGGATGTCATTTTGAAGAGGCGCAAAATCCCTTGTCTACGTA 771
Db 572 ---AAGTTGACCAAGCGCGCATGCTGGCAAAAGCGCAAAACCATGCTGTAAGT 627
Qy 772 GGGGAGGGGCGATGCGCCCATGCTCCAGTGCAGAGTGCAGGAATTTGCCGAAAGGTTG 831
Db 628 GGTGTGCGGTGGATATGCGCAGGAGTCTGCTTTACAGAAATTTCTGCTACACA 687
Qy 832 CAGTTGCGGTAAACAACCTGATGGAATTTGGGCTTTTGAAGAAACATTCCTT 891
Db 688 AAAATGCTGCCACCTGACCGCTGAAGAGGCTGGGCGGATGGAAGCATTAATCCGTAC 747
Qy 892 TCGGTGGATATGTTGGTATGATGACGACCGCTATGCAACTTTGCGGTGACGATGT 951
Db 748 TATCTGGGATGCTGGGAATGATGACCAAAAGCGGAACTTGCGGTGAGAGAGTGC 807
Qy 952 GATTGTTGATGACAGTGGGCGCCGTTTGAAGCAACCGGTTAATGTCGTAACGAA 1011
Db 808 GACTTCTGATCGCGTGGGTGACAGTTTGAATGACCGGTTGACCGGAAACTGAACACC 867
Qy 1012 TTTCGAGCGCGGCAAGTATTCATTTGACATGACACCGGCGGAGAGTGGGAAAAAC 1071
Db 868 TTGCGACACAGCGAGTGTATTCATATGATGACCGGCGGAAATGACAAAGCTG 927
Qy 1072 AAGGCTCCGATGTGCCATTTGTGGGAGATGATCCCATGTTTGAACAGCTTTGAG 1131

Db 928 CGTCAGGACATGTGGCATTTACAGATGATTTAAATGCTCTGTATCCAGCATTTACAGAG 987
 QY 1132 CGGGCCCCGGGAATTTGATTTACCCCAATCCCATACACAGGATGTTAAATGCG 1191
 Db 988 CCGTTAA-----ATATCAATGATGGCAGCTACACGCG 1020
 QY 1192 ATTGATCATTTGGCGGACCATTTACCCCTCCAGGTGCCCCACTATAGAGATTAATGCCC 1251
 Db 1021 GCGCAGCTGCGTGAATGAAATGCTGGGCTTACGACATCCCGGTGAGCGTATCTACGCG 1080
 QY 1252 CCCCAGAGGTAGTACAGAAATTTGTCGCGACGCCCCCGATGCTTACACACACGAT 1311
 Db 1081 CCATTGTTTAAACAATGCTGCGATGCTAACTGGGAGTTGGTGTGATCCACAGAT 1140
 QY 1312 GTGGGACAAACCAATGTGGGCGGCCAG--TTTGAACAATGGCCCCCGCGATGG 1368
 Db 1141 GTGGGACAGCACAGATGTGGGCGCGACACATCGCACACTCGCCGAAATTTTC 1200
 QY 1369 ATTCCAGTGTGCTGGGTGATGATGGGCTTTGGTTACCTGCGCGCATGGGAGCCAA 1428
 Db 1201 ATTACTTCAGCGGCTTTAGGACCATGGGTTTCGTTTACGACGCGGTTGGGCGACAA 1260
 QY 1429 GTGGAGTGGGGGACGA-GCGGTCAATTTCATCATGTGAGATGCGACGCTTCAATGAT 1487
 Db 1261 GTGCGACAGACGATGATGCTGTGTGTATCTCGGTGACGGCTTTTATGATGAT 1320
 QY 1488 CTTGAGAACTGGGGAACCTTACCCAGTACGACATCCAGTTAAACTTAATTTCTCAAT 1547
 Db 1321 GTGCAAGAGCTGGGACCGTAAACGCAAGCATTCAGTTAAATGCTTACTGAT 1380
 QY 1548 AACGTTGACAGGGAGTGTGCTGATGAGCAACAATTCTTCAAGAAAGCTTAATCT 1607
 Db 1381 AACCAAGGTTAGGATGTTGATGATGATGAGCAACGTTTTCAGGAACGATACAGC 1440
 QY 1608 GCTTTAAACATGTCCCAAGGACATGCAACATTAATCTCTGTGAAAGCTTATGCAATC 1667
 Db 1441 GAAACACCCCTTAC--TGATTAACCCCATTTCTCATGTTAGCCAGCGCTTGGGATC 1497
 QY 1668 AAGGTATTAATCTGTGCGAAGCGGAGATTTGGCCCCGGGAGATGCGGAATCTAGCC 1727
 Db 1488 CCGGACCAACATATACCCGTTAAAGACAGGTTGAAGGCACTGACACATGCTGAC 1557
 QY 1728 CACAATGTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1787
 Db 1558 AGTATGAGGACATACCTGCTTCAATGCAATGCAAGAACTTGAAGAGTGTGGCGGCTG 1617
 QY 1788 AATGGCCCCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
 Db 1618 GTGGCGCTGGCGCCAGTATTCAGAAATGTTGG 1651

RESULT 13
 US-10-096-571-13
 ; Sequence 13, Application US/10096571

; GENERAL INFORMATION:
 ; APPLICANT: RIEPING, MECHTHILD
 ; TITLE OF INVENTION: Process for the preparation of D-pantothenic acid and/or salts th
 ; FILE REFERENCE: 211499
 ; CURRENT FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: US/10/096,571
 ; PRIOR FILING DATE: 2001-05-14
 ; PRIOR APPLICATION NUMBER: US 60/304,776
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 2111
 ; TYPE: DNA

; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: RBS

; LOCATION: (8)..(12)
 ; OTHER INFORMATION:
 ; NAME/KEY: mutation
 ; LOCATION: (8)..(8)
 ; OTHER INFORMATION: Insertion of the base A at position 8
 ; NAME/KEY: CDS
 ; LOCATION: (1662)..(1925)
 ; OTHER INFORMATION:
 US-10-096-571-13

Query Match 16.3%; Score 318.8; DB 4; Length 2111;
 Best Local Similarity 52.8%; Pred. No.3.5e-95;
 Matches 873; Conservative 0; Mismatches 717; Indels 64; Gaps 6;

QY 1172 CATGGGCTCAACACATTTTGGCTATCCCGGCGGCAATTTGGCCATCTATGATGAA 231
 Db 58 CAGGATGTGAACACCGTTTCGGTTATCCGGGTGGCGCAATTAATGCGGTTTACATGCA 117
 QY 232 CTGTACCGCTTTGAACGCGGCGGGAATTTGACATATTTTGGTGGCCATGACAGCA 291
 Db 118 TTGTATGAC-----GGCGGCGTGAACCTTGTCTGTGCGGACATGACAGAGGT 165
 QY 292 GCTTCCCATGCGCGGATGGGTATGCGACAGCCACAGTAAAGTGGAGTTTGTTCGGT 351
 Db 166 GCGGCAATGGCGGCTATCGGTTATGCCCTGTCTACCGGCAAACTGGCTATGTATCGCC 225
 QY 352 ACATCTGACACAGGCGGCACTTAATTGTGATCGGCAATTTGCCATGCCATTTGACTG 411
 Db 226 ACCTGTGTGCGGCGGCAACCACTGATTAACCGGCTTGGCGGACCGCATGTTAGATTCT 285
 QY 412 GTGCCATGGTGTGATTAATCTGAGAGGTGGGCGGCGCATGATGATGATGATGATGATG 471
 Db 286 ATCCCTGTTGTTCCATACCGGTCAAGTGTGCGCACCGTTTATCGGACAGAGCATTT 345
 QY 472 CAGGAATTTGACATTTTGGCATCACCTTAACCATGTTAAGCATCTTATGTGTACGT 531
 Db 346 CAGGAATGATGATCTCGGATTTGCTGTTAGCTGTACCAAGACACGCTTTCTGTGACG 405
 QY 532 AGTGGCGGATATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
 Db 406 TCGCTGAAGAGTTGGCGGCACTTAATGCTGAAGCAATTTGACGTTGCCAGCTCAAGTGTCT 465
 QY 592 CCGGCGCGGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 651
 Db 466 CTGTGTGCGGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 525
 QY 652 CCCCTGACCCCGGTACGTTAATCTAACCGGTTATGCGCCACGCTTAAAGTAAATCCC 711
 Db 526 CCGTGTGATCAACCCGTTGAAACGAAGTACTTCCCATGCGC-----571
 QY 712 CGACAATTAATCGGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 771
 Db 572 ----AACTTGAAGACAGCGGCGGCAATGCTGCGCAAAAGCCCAAAACGATGCTATCGTT 627
 QY 772 GGGGAGGGGCGATGCGCCGCAATGCGCATGCCAGTGCAGAAATTTGCGGAAAGTTCC 831
 Db 628 GGTGTGTGGGTGTGATGCGGCGAGGAGTTCGTTTACGAAATTTCTCCTACACACA 687
 QY 832 CATTTCCGCTTACCAACCCGATGAGGAAATTTGGGCTTTTGAAGAAACATCCCTT 891
 Db 688 AAATGCTGTCACTGCGACGCTGAAGAGGCTGGGCGGCTTGAACCAATTAATCCGTAC 747
 QY 892 TCGGTGGTATGTTGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 951
 Db 748 TATCTGGGATGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 807
 QY 952 GATTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1011
 Db 808 GACTTGTATGCTGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 867
 QY 1012 TTGTAGCCGCGCCAAATTAATTCATTTGATGATGATGATGATGATGATGATGATGATG 1071
 Db 868 TTGTAGCCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927

1072 AGGGCTCCCGATGTCCTATGTCGGGATGTCAGCCATGTTTGAACAGCTTTTGAG 1131
928 COTAGGACATGTCGATTAACAAGGTCATTAATGTCGTGTACACAGATTCACAG 987
1132 CGGGCCCGGAAATTGATTAACCCACCATCCCATACACCCAGGATGTTAAATCCG 1191
988 CCGTTAA-----ATATCATGTCGCGACCTACCTGC 1020
1192 ATTGATCATGTCGCGACCATTAACCCCTCCAGGTGCCCCACTATGAGATTAATGCC 1251
1021 GCGAGAGTCGCTGATGAACATGCTGGGTTACACACATCCCGGTGACGCTATGACGC 1080
1252 CCCCAGAGGTATGACAGAAATTGTCGCCAGGCCCGCCGATGCTTCTACACCCAGAT 1311
1081 CCATGTTGTTAAACACATGTCGATTCGTAACCTCGGATTCGCTGTCGACACAGAT 1140
1312 GTGGACACACACCAATGTCGGCGCCAG---TTTTGAACAATGCCCCCGCCAGATG 1368
1141 GTGGGCGACACAGATGTCGGCGCCGAGACATCCGACACACTCCCGGAAAAATTC 1200
1369 ATTTCAGTGTGCTGCTGGTGGTACATGCGCTTTGCTTACCTGCGCCATGAGCCAA 1428
1201 ATTACTCTCAGCGGCTTAGGACACATGGGTTTCGTTTACGACGCGGCTTGCGCACAA 1260
1429 GTGGGAGTGGGGGACGA--GGGTCAATTGATGATGAGATGCGACCTTCCAAATGAT 1487
1261 GTCCGACGACCGAAGCATCTGTCGTCTGTATCTCCGCTGACGCTCTTTCATGATGAT 1320
1488 CTTGAGAACTGGGACCCCTAGCCAGTACGACATCCAGGTTAAACTATATTTCTCAAT 1547
1321 GTGCAAGAGTGGGACCGTAAACGACAGATTAACGTTGAAATACTGCTTACTGAT 1380
1548 AACGGTTGGCAGGGATGTCGTCAGTGCAGCAAACTTCTACGAGAAAGCTATCT 1607
1381 AACCAAGCTTAAAGGATGTCGACAAATGSCAGCACTGTTTTCAGAAACGATAGC 1440
1608 GCTTCTAATCATGTCGACGAGCATCCAGACATTAATCTCTCTGTAAGCTATGAGCAT 1667
1441 GAACCACTCTTAC---TGATTAACCCGATTTCTCATGTTAGCAGCGCTTGCGGATC 1497
1668 AAGGTATTAATGTCGCGACGCGGAGATTTGGCCCCGCGATGCGGAAATGCTAGCC 1727
1498 CCGGCAACACATCAACCCGTAAAGACCGGTTAAGCGGACCTCGACACATGCTGATC 1557
1728 CACATGTCCTGTCGTCATGATGATGTCGTAAGAAAGATGAAACTGTTACCTATG 1787
1558 AGTATGAGGACATACCTCTCTCATGTCTCATTCACATCGACGAACTGGAACGTCGCGCTG 1617
1788 ATTGCCCCCGCATGATGATGTCGCAATGCTAG 1821
1618 GTGCGCTGCGCGCAGTAAATTCAGAAATGTTG 1651

RESULT 14
US-08-452-075-1
Sequence 1, Application US/08452075
Patent No. 5998178
GENERAL INFORMATION:
APPLICANT: HASHIGUCHI, KEN-ICHI
APPLICANT: KISHINO, HIROKO
APPLICANT: TSUTSUMOTO, NOBUHARU
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: L-ISOLEUCINE-PRODUCING BACTERIUM AND
METHOD FOR PREPARING L-ISOLEUCINE THROUGH FERMENTATION
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,075
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 116340/1994
FILING DATE: 30-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 08/452,075
REFERENCE/DOCKET NUMBER: 10-741-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2841 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 957..1052
FEATURE:
NAME/KEY: CDS
LOCATION: 1195..2838
US-08-452-075-1

Query Match 15.8%; Score 309.2; DB 2; Length 2841;
Best Local Similarity 52.4%; Pred. No. 6.7e-92;
Matches 867; Conservative 0; Mismatches 723; Indels 64; Gaps 6;

172 CATGGGTCACAAACATATTTTGGCTATCCCGCGGGGCAATTTTGGCCATCTATGATGA 231
1234 CAGGTGTGACACCGCTTTTGGCTATCCGGGTGGCGCAATTATCCCGTTTACATGCA 1293
232 CTGTAACCGTTTGAACGCGCGGGGAAATTGACATATTTTGGTCGCCATGAAACAGCA 291
1294 TTGTATGAC-----GGCGCGCTGAGACCTTGCTATGCGGACATGAGCAGGCT 1341
292 GCTTCCCATGCGCGGATGGGTATGCGCAGACGCAAGGTTAAAGTGGAGTTGTTGGGT 351
1342 GCGGCATAGCGCGGCTATGCGGTATGCTGCTACCGGCAAACTGCGCTATGATGCGC 1401
352 ACATGTGACACAGGCGCATTAATCTGTGACACCGCATTTGCCAATGCCATTGGACTCG 411
1402 ACGTCTGGGCCGCGGCAACCACTGATAACCGGCGCTTGGGAGCGCACTGTGATTC 1461
412 GTGCCATGTCGTATTAATGAGAGAGTGGGCGCTGCAATGATGTTGATGAGATGCTTTC 471
1462 ATCCCTGTTTGGCATACCGGTCAAGTGTCCGACCGTTTATGCGCATGACGATTT 1521
472 CAGAAATGACATTTTGGCATCACTTACCGATCGTTTAAAGCATCCCTATGTCGTAGT 531
1522 CAGAAAGTGAATGCTCTGGAAATGCTGCTTACCTGTACCAAGCATGATCTTCTGTCAG 1581
532 AATGCGCGGATATGCTGCTGCAATTTTACTGAGGCTTTCATCTTCTGACAGCGGTGT 591
1582 TCGCTGGAAGATTCGCCCGCATCAATGCTTAAGATTCAGCTGCTGCTCAGGTGT 1641
592 CCGGCGCGGTTTGTATGATGATTCACCAAGATGAGGCTTAGAAGATGAGATCAAT 651
1642 CCGGTCCGCTTCTGTGATATCCCAAAATATCAGTTAGCCAGGCGTGAACCTGGA 1701
652 CCGCTGACCCCGGTGACGTATATCTACCGGCTTATGCGCCGCTTAAAGTATATCC 711

Db 1702 CCGTGGTTCACCAACCGTGGAAAAAGAGTACTTTCCCATGCGG----- 1747
Qy 712 CGACAATTAATATGCGGCATTTGCAATTTGGAGCAGGCGAGAAATCCCTGCTCACTGA 771
Db 1748 ----AAGTTGAGCAAGCGCGCCAGATGCTGGCAAAAGGCAAAACCAATGCTGACCTT 1803
Qy 772 GGGGGAAGGGGCGATGCGCGCCATGCGCATGCGCAAGTGGCGAATTTGCGGAAAGGTC 831
Db 1804 GCGCGATGGGTGGGTATGCGCGAGGCAAGTTCGCGCTTGGCGGAATTTCTGCTGCCACA 1863
Qy 832 CAGTGGCGGTACACCAACCGTGAATGGGAATTTGGGGCTTTTGAAGAAACCATCCCTT 891
Db 1864 AAAATGCTGCTGCACTGTAAGCTGAAGAGGCTGGCGCAGTAAGAGCAATTAATCCGAC 1923
Qy 892 TCGGTGGGTATGTTGGGTATGATGAGCAACCGCTATGCGCACTTTGCGGTGACGCGAATGT 951
Db 1924 TATCTGGGCGATGCTGGGGATGCAACGACCAAAAGCGCAAACTTGGCGGTGAGAGAGTGT 1983
Qy 952 GATTGTTGATGCAATGGGGGCGCGTTTCCAGCGACCGGGTAACTGGCAAACTAGACGA 1011
Db 1984 GACCTGCTGATGCGGTGGGCGCAGTTTGTATGACCGGTGACCGGCAAACTGAACACC 2043
Qy 1012 TTTGCTAGCGCGCCCAAAATTAATTCATTCATGACATGACCGCGGAGGTGGGAAAAAAC 1071
Db 2044 TTGCGGCGCACACGCGCAGTTTATTCATATGATATGACCGCGGAGAAATGACAAAGCTG 2103
Qy 1072 AGGGTCCCGATGTGCCCATTTGTGGGGATGTAGCCCATGTTTGAACAGCTTTTGCGAG 1131
Db 2104 CGTACAGCACAATGTGCAATTCACAAAGTATTAATATGCTCTGTACCGCATTCACAGAG 2163
Qy 1132 CGGGCCCGGGAATTTGATTAACCCCAACCATCCCATACACCGAGCATGTTAAATGCG 1191
Db 2164 CGCTTAATCAATGTGACTGCGA-----GCAACACTGCG 2196
Qy 1192 ATTGATCATTGGCGGACCGATTACCCCTCCAGGTGCCCACTATAGAGATTAATGCGC 1251
Db 2197 GCGCAGCTGCGTGAATGAAATCTCTGGCGTACGACATCCCGGTGACGCTATCTAGCGG 2256
Qy 1252 CCCCAGAGGTAGTACAGAAATTTGTGCGCAGCGCCCGCATGCTTAACACACCGAT 1311
Db 2257 CCGTGTGTTTAAACAATGCTCGGATGTAACTTGCGGATGTGTGACACACAGAT 2316
Qy 1312 GTGGAGCAACACCAATGTGGCGCGCCAG--TTTGAACAATGCGCCCGCGCGATG 1368
Db 2317 GTGGGGCAGACCAAGATGGGCTGCGAGCAGCATGCGCCACACTGCGCGGAAATTTTC 2376
Qy 1369 ATTTCAGTGTGCTGTGGGTACGATGGGCTTTGTTTACTGCCCGCATGGAGCCAAA 1428
Db 2377 ATCACCTCCAGCGGTTTAAAGTAAACATGGGTTTGGTTTACCGGCGCGGTGGGCGCA 2436
Qy 1429 GTGGAGTGGGGAGCA--GCGGTCAATTGACATGAGATGAGCATTCCTCAATGAT 1487
Db 2437 GTCGGGCGCAGCAAGATACCTGTTGTAATCTCGGTGACGAGCTCTTTATGATGAT 2486
Qy 1488 CTTCAAGAACTGGGAACCTTGAACCCAGTACGACATCCAGTTTAAACTATTAATTTCTCAAT 1547
Db 2497 GTGCAAGAGCTGGGCAACGTAAACGCAAGAGATTCGTTGAATGCTTTACTCGAT 2556
Qy 1548 AACGGTTGGAGGGATGTTGCGTGAATGGGCAACAACTTTCTAGAAAGATTAATCT 1607
Db 2557 AACCAAGGTATGAGATGTTGACAAATGGCAGCACTGTTTCTCAAGAAAGATTAACGC 2616
Qy 1608 GCTTCAACATGTCCAGGGATGCGCAACATTAATCTCTCTGTGAAGCCTATGCGATC 1667
Db 2617 GAAACCAACCCCTTAC--TGATTAACCCGATTTCTCATGTATGACAGGCGCTTGGGATC 2673
Qy 1668 AAGGGTATTAATCTGTGCGCAAGCGGGAAGATTTGGCCCGGGGATGCGCGAATCTAGCC 1727
Db 2674 CATGGCCAAACATACCCCGAAGCAAGGTTGAAGCGGCACTGGAACCATGCTGAAC 2733
Qy 1728 CACAAATGCTCTGTGTATGATGATGTGTGTCAAAAAGATGAAGAACTGTAACTATG 1787
Db 2734 AGTATGGGCGCATACCTGCTTCAATGTCATGCAAGAACTTGAAGAGTGTGGCGCGTG 2793

Qy 1788 ATTGCCCCCGCATGATGATATGCCCCAAATGCTAG 1821
Db 2794 GTGCGCGCTGCGCCAGTATTAATGAAATGTTGG 2827

RESULT 15
US-09-231-061-1
: Sequence 1, Application US/09231061
: Patent No. 6214591
: GENERAL INFORMATION:
: APPLICANT: TOMITA, FUSAO
: APPLICANT: YOKOTA, ATSUSHI
: APPLICANT: HASHIGUCHI, KENICHI
: APPLICANT: ISHIGOOKA, MASAKO
: APPLICANT: KURAHASHI, OSAMU
: TITLE OF INVENTION: METHODS FOR PRODUCING L-VALINE AND
: TITLE OF INVENTION: L-LEUCINE
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSER: P.C.
: STREET: 1755 S. JEFF. DAVIS HIGHWAY, FOURTH FLOOR
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/231,061
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/793,441
: FILING DATE: 28-FEB-1997
: APPLICATION NUMBER: PCT/JP95/01719
: FILING DATE: 30-AUG-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-204856
: FILING DATE: 30-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 10-843-0 PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220
: INFORMATION FOR SEQ. ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2841 base pairs
: type: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: ESCHERICHIA COLI
: STRAIN: M1162
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 957..1055
: OTHER INFORMATION: /note= "IDENTIFICATION METHOD: S"
: FEATURE:
: NAME/KEY: attenuator
: LOCATION: 1081..1104
: OTHER INFORMATION: /note= "IDENTIFICATION METHOD: S"
: FEATURE:
: NAME/KEY: CDS

```

/ LOCATION: 1195..2841
/ OTHER INFORMATION: /note= "IDENTIFICATION METHOD: S"
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 52..57
/ OTHER INFORMATION: /note= "CLEAVAGE SITE: SmaI,
/ OTHER INFORMATION: IDENTIFICATION METHOD: S"
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 2395..2400
/ OTHER INFORMATION: /note= "CLEAVAGE SITE: KpnI,
/ OTHER INFORMATION: IDENTIFICATION METHOD: S"
US-09-231-061-1

Query Match      15.8%; Score 309.2; DB 3; Length 2841;
Best Local Similarity 52.4%; Pred. No. 6.7e-92;
Matches 867; Conservative 0; Mismatches 723; Indels 64; Gaps 6;

QY 172 CATGGGCGTCAACACATTTTGGCGTATCCCGGCGGCGCAATTTGGCCATCTATGATGAA 231
DB 1234 CAGGGTGTGAACACCGTTTTCGTTATCCGGGTGCGCATTTATGCGGTTTACGATGCA 1293
QY 232 CTGTACCGCTTTGAAGCGCGGCGGAAATTGAGCATATTGTGCGCCATGAA 291
DB 1294 TTGTATGAC-----GGCGGCGTGAAGCACTTGCTATGCCGACATGAGCAGGCT 1341
QY 292 GCTTCCCATGGCGGAGATGGGTATGCCAGACCCACAGTAAAGTGGAGTTGTTCGT 351
DB 1342 GCGGCAATGGCGGCTATCGTTATGCTGCTACCGGCAAACTGGCGTATGATGCC 1401
QY 352 ACATCTGACACGCGGCGCACTACTGCTGTGACCGGCACTTGGCAATGCCATTTGSACTCG 411
DB 1402 ACGTGTGTCGGGCGGCAACCACTGATTAACCGGGCTGGGAGCGACATGTTGATTC 1461
QY 412 GTGCCCATGTGTGTGATTACTGAGAGGTGGCGGTCGCATGATGTGTAGCATGCTTTC 471
DB 1462 ATCCCTGTGTGTGCTACCGGCTCAAGTGTCCGACCGTTTACCGGACCTGAGCAATTT 1521
QY 472 CAGGAATTTGACATTTTGGCATCACCTTACCGATCGTTAGACCTCTGATGATGCT 531
DB 1522 CAGGAATGTGATGTCCTGTGATGTGCTTACCTGTACCAAGCATACCTTTTGTGTCAG 1581
QY 532 AGTGGCGGATATGCTGCAATTTGTAATGAGGCTTTTCCATCTTGTGACACCGGTCGT 591
DB 1582 TCGCTGGAAGATGTCGCCGATCATGCTGATGACATTCGACGTGCTGTCAGGTCGT 1641
QY 592 CCCGCGCGGTTTGTGATGATGATTCGCAAGGATGCGGCTTAGAAGATGTGATACAT 651
DB 1642 CCGGTGCGGTTCTGTGATGATTCGCAAGGATGCAATGATGACGAGGTTGACCTGAA 1701
QY 652 CCCCTCGACCCCGGTGACGTTATCTACCGGTTATGCGCCACGCTTAAAGTATCCC 711
DB 1702 CCGTGTTCACCAACCGTTGAAAGCAAGTGAATTTCCCAATGCG-----1747
QY 712 CGACAAATTAATGGCGCATTTGCAATTTGTGAGAGAGCCAGAAATCCCTTGTCTAGCTA 771
DB 1748 -----AAGTTGAGCAAGCGGCAAGATGCTGCAAAAGGCAAAACCGATGCTGTATGCTT 1803
QY 772 GGGGAGGCGGATGCGCGCCATGCGCCATGCCAGGTGCGAGATTTGGGAAAGGTTTC 831
DB 1804 GCGGCTGCGGTGGATGCGGCAAGCAATTCGCGCTTTCGCTGATTTTCGCTGCACACA 1863
QY 832 CAGTTTCCGCTTAACACACCTCTGATGGAATTGGGCTTTTGAAGAAACCATCCCTT 891
DB 1864 AAAATGCTGCGCACTGTACCTGAAAGGCTGGGCGAGTAGAAGCAGATTAATCCGTAC 1923
QY 892 TCGGTGGGTATGTTGGGTATGCAATGCGCATATGCACTTGTGCGTCAAGGATGT 951
DB 1924 TATCTGGGATGCTGGGATGCAAGCAACAAAGCGCAAACTTTCGCGGTGCAAGATGT 1983
QY 952 GATTTGTTGATGCAATGCGGCGCTTTCGACGACCGGCTAACTGCGCAAACTAGAGAA 1011
DB 1984 GACCTGTGATGCGCGTGGCGCAGTTTGTATGACCGGCTGACCGGCAAACTGAAACACC 2043

```

```

QY 1012 TTGTGATCCGCGCCAAAGTAAATTCATATGACATGACCCGCGGAGATGGGAAAAAAC 1071
DB 2044 TTGCGGCCACAGCCAGTGTATTCATATGATATGACCCGCGGAGAAATGAAACAGCTG 2103
QY 1072 AAGGCTCCGATGTCCTCATTTGTGGGATGTAACCCATGTTTAAAGACACTTTTGAG 1131
DB 2104 CGTCAAGGACATGTGTGATTAACAGATGTTAAATGCTGTATCCGACATTAACAGAG 2163
QY 1132 CGGCGCCCGGAAATTTGATTAACCCACCATCCCATTAACACCCAGGATGTTAAATGC 1191
DB 2164 CCGTTAAATCAATGTGACTGGCA-----GCAACACTGC 2196
QY 1192 ATTGATCATTTGGCGGACCGATTAACCCCTCCAGTGGCCCACTATGAGATATATTC 1251
DB 2197 GCGCAGCTGCGGATGAACATTTCTGGGTTACGACATCCCGGTGACGCTATCTACGCG 2256
QY 1252 CCCGAGAGGTAGTACAGAAATTTGTGCGCAGGCGCCGATGCTTACATACACACCGAT 1311
DB 2257 CCGTTGTTTAAACATGTGTGATGATGTAACCTGCGGATGTGCTGTACACACAGAT 2316
QY 1312 GTGGGACAAACACCAATGTGGCGGCGCAG---TTTGAACATGCGCCCGCGATGG 1368
DB 2317 GTGGGCGAGCACAGATGTGGGCTGCGGACACATGCGCCACACTGCGCGGAAATTTTC 2376
QY 1369 ATTGCAAGTGTGGCTGGGTGATGAGTGGGCTTTGATACCTGCGCGCATGAGGACCAA 1428
DB 2377 ATCACTCCAGCGCTTTAGTATGATGATGATGATGATGATGATGATGATGATGAT 2436
QY 1429 GTGGAGATGGGAGGAGCA--GCGGTCAATTTGATCACTGATGATGATGATGATGAT 1487
DB 2437 GTGCGGCGACCGAAGATATCCGTGTGTATCTCGGTGACGCGCTTTTATGATGAT 2496
QY 1488 CTTGAGGAATGAGGAACCTTACCCAGTACGACATCCAGGTTAAACATATATTCAT 1547
DB 2497 GTGCAAGTGTGGGACCGTAAACGCAAGCTTACCTGTTAAATGCTCTTACTCAT 2556
QY 1548 AACGTTGGCAAGGATGTGGTGTGATGCAAGCAAACTTTCTAGAAAGAGCTTATCT 1607
DB 2557 AACCAACGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2616
QY 1608 GCTTGAACATGTCCCAAGGACATGCGACATTAATCTCTGTGAAAGCTTATGCGCAT 1667
DB 2617 GAACCAACCTTAC---TGATTAACCCGATTTCTCATGTTAGCCAGGCGCTTGCGCATC 2673
QY 1668 AAGGATTAATCTGTGCGCAAGGAGGAAATTTGGCCCGGATGCGCGAAATGCTACCC 1727
DB 2674 CATGGCCAAACATCAACCGGAAAGACAGGTGAAGGCGCATGTGACATGCTGAAC 2733
QY 1728 CACAAATGCTGTGTGTGATGATGTGTGTGATGATGATGATGATGATGATGATGAT 1787
DB 2734 AGTGAATGGGCAATACCTGCTTCATGTCTCAATGCAAGCAATGGAACGTCGTGGCCGCTG 2793
QY 1788 ATTGCCCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1821
DB 2794 GTGCGGCTGCGCGCAATGATTAATCAGAAATGTTGG 2827

```

Search completed: July 25, 2004, 12:54:47
 Job time : 166.901 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 07:28:18 ; Search time 943.797 Seconds

(without alignments)
10146.437 Million cell updates/sec

Title: US-09-893-033-6

Perfect score: 1955

Sequence: 1 gccatagagaccatcgccg.....cataagccaaatgcaatc 1959

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3216467 segs, 244149694 residues

Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	1843.4	94.1	1865	US-10-369-493-26409	Sequence 26409, A
2	847.2	43.2	1815	US-10-369-493-43721	Sequence 43721, A
3	841.8	43.0	1815	US-10-369-493-42620	Sequence 42620, A
4	415	21.5	1728	US-10-282-122A-41727	Sequence 41727, A
5	381.6	19.2	1878	US-10-282-122A-21733	Sequence 21733, A
6	379.8	19.4	1866	US-10-282-122A-25560	Sequence 25560, A
7	379.6	19.4	1722	US-10-282-122A-41305	Sequence 41305, A
8	378.8	19.3	1875	US-10-282-122A-11812	Sequence 11812, A
9	373.8	19.1	1725	US-09-815-242-7947	Sequence 7947, Ap
10	373.8	19.1	1725	US-10-282-122A-30524	Sequence 30524, A
11	368.6	18.8	1815	US-09-815-242-5925	Sequence 5925, Ap
12	368.6	18.8	1815	US-10-282-122A-20222	Sequence 20222, A
13	368.2	18.8	1833	US-10-282-122A-26894	Sequence 26894, A
14	368.2	18.8	96109	US-10-672-787-35	Sequence 35, Appl

15	367.8	18.8	1725	13	US-10-282-122A-30934	Sequence 30934, A
16	366.8	18.7	1854	13	US-10-282-122A-28264	Sequence 28264, A
17	366.8	18.7	1857	13	US-10-282-122A-28639	Sequence 28639, A
18	363	18.5	1596	13	US-10-282-122A-32093	Sequence 32093, A
19	363	18.5	1725	17	US-10-875-026A-29	Sequence 29, Appl
20	363	18.5	1725	17	US-10-875-026A-95	Sequence 95, Appl
21	363	18.5	1728	13	US-10-282-122A-29827	Sequence 29827, A
22	358.6	18.3	1782	16	US-10-369-493-35567	Sequence 35567, A
23	357	18.2	1848	15	US-10-156-761-2720	Sequence 2720, Ap
24	357	18.2	9025608	15	US-10-156-761-1	Sequence 1, Appl
25	355	18.1	1725	13	US-10-282-122A-29151	Sequence 29151, A
26	354	18.1	1722	13	US-10-282-122A-24462	Sequence 24462, A
27	350.8	17.9	6506	16	US-10-398-221-3845	Sequence 3845, Ap
28	349.8	17.9	1564	13	US-10-282-122A-13402	Sequence 13402, A
29	345.8	17.7	1554	13	US-10-282-122A-13627	Sequence 13627, A
30	345	17.6	2286	16	US-10-398-221-1981	Sequence 1981, Ap
31	345	17.6	319630	16	US-10-398-221-7	Sequence 7, Appl
32	345	17.6	3011208	16	US-10-398-221-2058	Sequence 2058, Ap
33	342.2	17.5	1713	13	US-10-282-122A-15103	Sequence 15103, A
34	337.8	17.2	1956	13	US-10-282-122A-17528	Sequence 17528, A
35	337.2	17.2	1761	13	US-10-282-122A-14449	Sequence 14449, A
36	330.2	16.9	1642	13	US-10-282-122A-19815	Sequence 19815, A
37	329.4	16.8	1722	9	US-09-815-242-7160	Sequence 7160, Ap
38	329.4	16.8	1722	13	US-10-282-122A-22365	Sequence 22365, A
39	329.2	16.8	1878	9	US-09-738-626-3457	Sequence 3457, Ap
40	329.2	16.8	3309400	9	US-09-738-626-1	Sequence 1, Appl
41	328.4	16.8	1830121	15	US-10-329-960-1	Sequence 1, Appl
42	328.4	16.8	1830121	16	US-10-329-960-1	Sequence 1, Appl
43	325.4	16.6	1644	13	US-10-282-122A-23451	Sequence 23451, A
44	323.4	16.5	1722	13	US-10-282-122A-9061	Sequence 9061, Ap
45	321.4	16.4	1701	13	US-10-282-122A-36338	Sequence 36338, A

ALIGNMENTS

RESULT 1
US-10-369-493-26409
Sequence 26409, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 26409
LENGTH: 1866
TYPE: DNA
ORGANISM: Synechocystis sp.
US-10-369-493-26409

Query Match 94.1%; Score 1843.4; DB 16; Length 1866;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1858; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

73 GGTGTGCATCAATTCCTCAATCCCAAAACAATTTCTGAAACTGTTCTAGCAACGG 132
1 GTGGTTCATCAATTCCTCAATCCCAAAACAATTTCTGAAACTGTTCTAGCAACGG 60

133 CAACCGGGGCTATATCTCGATGATAGCTGAACGCGATGGGTCAACACATTTT 192
61 CAACCGGGGCTATATCTCGATGATAGCTGAACGCGATGGGTCAACACATTTT 120
193 GGTATCCCGGCGGGGCAATTTTGGCCATCTATGATGAACCTGACCGCTTTGAAGCGGG 252

```

Db      121 GGTATCCCGCGGGCAATTTTSCCATCTATGATGAACCTGTAACCCCTTGAAGCGGC 180
Qy      253 GGGGAAATGAGCATTTTGTGTGCGCATGAAGAGAGTCCCATGCGGGGANTGG 312
Db      181 GGGGAAATGAGCATTTTGTGTGCGCATGAAGAGAGTCCCATGCGGGGANTGG 240
Qy      313 TATGCGAGAGCCAGGTAAGTGGAGTTTGTGCGTACATCTGGAACGAGGCGACT 372
Db      241 TATGCGAGAGCCAGGTAAGTGGAGTTTGTGCGTACATCTGGAACGAGGCGACT 300
Qy      373 AACTTGTGAGCGGCATTTGCGCATTTGGAACCTGCGGCGCATGAGTGTACTACT 432
Db      301 AACTTGTGAGCGGCATTTGCGCATTTGGAACCTGCGGCGCATGAGTGTACTACT 360
Qy      433 GGAAGGTGCGCGCTGCGCATTTGTAAGCATCTTTCAGAGAAATGACATTTTGGC 492
Db      361 GGAAGGTGCGCGCTGCGCATTTGTAAGCATCTTTCAGAGAAATGACATTTTGGC 420
Qy      493 ATCACTTACCGATGCTTGAAGCATCTCTATGTGTAGTGTGCGGAGATAGGCTCGC 552
Db      421 ATCACTTACCGATGCTTGAAGCATCTCTATGTGTAGTGTGCGGAGATAGGCTCGC 480
Qy      553 ATTGTTACTGAGGCTTTCATCTTGTAGCAACCGGTGTCGCGGCGGTTTGTGATGAT 612
Db      481 ATTGTTACTGAGGCTTTCATCTTGTAGCAACCGGTGTCGCGGCGGTTTGTGATGAT 540
Qy      613 ATTCCCAAGATGTGGCTTGAAGATGTAGTACATTCCTCCGACCCCGGTGACGTT 672
Db      541 ATTCCCAAGATGTGGCTTGAAGATGTAGTACATTCCTCCGACCCCGGTGACGTT 600
Qy      673 AATCTACCGGGTTATGCGCCGCAAGGTTAAAGTAACTCCCGACAAATTAAGCGGCAATG 732
Db      601 AATCTACCGGGTTATGCGCCGCAAGGTTAAAGTAACTCCCGACAAATTAAGCGGCAATG 660
Qy      733 CAATGTTGAGCAAGGCGCAAAATCCCTTGTCTACGTAGGAGGAGGCGATGCGCGC 792
Db      661 CAATGTTGAGCAAGGCGCAAAATCCCTTGTCTACGTAGGAGGAGGCGATGCGCGC 720
Qy      793 AATGCGCATGCGCGAGTGCAGAAATTTGCGGAAGGTTCCAGTTGCGGTAACAACACC 852
Db      721 AATGCGCATGCGCGAGTGCAGAAATTTGCGGAAGGTTCCAGTTGCGGTAACAACACC 780
Qy      853 CTGATGGGAATTGGGGCTTTTGAAGAAACCATCCCTTGTGAGTGTGGTATG 912
Db      781 CTGATGGGAATTGGGGCTTTTGAAGAAACCATCCCTTGTGAGTGTGGTATG 840
Qy      913 CATGCGCACCGCTATGCCAATCTTTCGCTGCGCATGATGTTTGTGATGCGAGTGGG 972
Db      841 CATGCGCACCGCTATGCCAATCTTTCGCTGCGCATGATGTTTGTGATGCGAGTGGG 900
Qy      973 GCCCGTTTGAAGCAAGGCGTAACTGAGCAATTTGCTAGCCGCGCAAGTA 1032
Db      901 GCCCGTTTGAAGCAAGGCGTAACTGAGCAATTTGCTAGCCGCGCAAGTA 960
Qy      1033 ATTCAATTGACATGACCGCGGAGTGGGAAAAAAGAGGCTCCCGATGTGCCATT 1092
Db      961 ATTCAATTGACATGACCGCGGAGTGGGAAAAAAGAGGCTCCCGATGTGCCATT 1020
Qy      1093 GTGGGGGATGATACCGCATGTTTGAAGCACTTTTGAAGCGGCGCGGAAATGGATTAC 1152
Db      1021 GTGGGGGATGATACCGCATGTTTGAAGCACTTTTGAAGCGGCGCGGAAATGGATTAC 1080
Qy      1153 CCCACCATCCCATACCAACCGAGCATGTTAAATGCGATGATGATGCGGACCGAT 1212
Db      1081 CCCACCATCCCATACCAACCGAGCATGTTAAATGCGATGATGATGCGGACCGAT 1140
Qy      1213 TACCCCTTCAGAGTCCCATATGAGATATCTATGCCCCCGAGAGTATGTAACAGAA 1272
Db      1141 TACCCCTTCAGAGTCCCATATGAGATATCTATGCCCCCGAGAGTATGTAACAGAA 1200
Qy      1273 ATTGGTCCGACAGGCCCGCATGCTACTACACCAACCGATGTGGACAAACCAATGTGG 1332

```

```

Db      1201 ATTGTGCGCAGGCCCCCGATGCTACTACACCAACCGATGTGGACAAACCAATGTGG 1260
Qy      1333 GCGGCCAGTTTGTGAAATGAGCGCCCGCGATGATTTTCAGTGTGCTTGGGTAAG 1392
Db      1261 GCGGCCAGTTTGTGAAATGAGCGCCCGCGATGATTTTCAGTGTGCTTGGGTAAG 1320
Qy      1393 ATGGGCTTTGTTTACCTGCGCGCATGAGGACCAAAAGTGGAGTGGGGGAG-AGCGGTC 1451
Db      1321 ATGGGCTTTGTTTACCTGCGCGCATGAGGACCAAAAGTGGAGTGGGGGAGCGGTC 1380
Qy      1452 ATTGTCATCAGTGGAGTCCAGGCTTCCAAATGATCTTCAGAACTGGGAACCTTAGCC 1511
Db      1381 ATTGTCATCAGTGGAGTCCAGGCTTCCAAATGATCTTCAGAACTGGGAACCTTAGCC 1440
Qy      1512 CAGTACGATCAGGTTTAAATCTATTTCTCAATTAACGTTGGCAGGGGATGTGCT 1571
Db      1441 CAGTACGATCAGGTTTAAATCTATTTCTCAATTAACGTTGGCAGGGGATGTGCT 1500
Qy      1572 CAGTGGCAACAACTTTCTACGAGAAAGTATTTCTGCTTCTTAACATGTCCAGGGCATG 1631
Db      1501 CAGTGGCAACAACTTTCTACGAGAAAGTATTTCTGCTTCTTAACATGTCCAGGGCATG 1560
Qy      1632 CCAGACATTAATCTCTCTGTGAAGCTATGSCATCAAGGTAATTAATCTGTGCGCAAGCG 1691
Db      1561 CCAGACATTAATCTCTCTGTGAAGCTATGSCATCAAGGTAATTAATCTGTGCGCAAGCG 1620
Qy      1692 GAAATTTGGCCCCGCGCATGCGGAAATGCTAGGCCAATGTCCTGTGGTATGAT 1751
Db      1621 GAAATTTGGCCCCGCGCATGCGGAAATGCTAGGCCAATGTCCTGTGGTATGAT 1680
Qy      1752 GTGTGTGTCAAAAAAGATGAATACTGTATCCCTATGATGTGCCCCGCGCATGATGCC 1811
Db      1681 GTGTGTGTCAAAAAAGATGAATACTGTATCCCTATGATGTGCCCCGCGCATGATGCC 1740
Qy      1812 CAAATGCTAGTTTACCGGAATGCGGTAACNGACATGTGTCCCGGATGTTGAGAGTC 1871
Db      1741 CAAATGCTAGTTTACCGGAATGCGGTAACNGACATGTGTCCCGGATGTTGAGAGTC 1800
Qy      1872 AACCATGTGCAACCAAAATTTCTATCAACCATGTTTCTGTGTTGTGAGCCAAA 1931
Db      1801 AACCATGTGCAACCAAAATTTCTATCAACCATGTTTCTGTGTTGTGAGCCAAA 1860
Qy      1932 CTCTAA 1937
Db      1861 CTCTAA 1866

RESULT 2
US-10-369-493-43721
/ Sequence 43721, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianheng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ NUMBER OF SEQ ID NOS: 43721
/ SEQ ID NO 43721
/ LENGTH: 1815
/ TYPE: DNA
/ ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-43721

Query Match 43.2%; Score 847.2; DB 16; Length 1815;
Best Local Similarity 67.5%; Pred. No. 1,3e-273;

```

Matches 1221; Conservative 0; Mismatches 584; Indels 4; Gaps 2;

```
QY 126 CCACAGGCAAAACGGGGCTTATATCTGATGATAGCTGAAGCGCCATGGGGTCAAA 185
Db 9 CAAAGCTGATCTGGGGTTTGTCTTACTTGAAGCTTCAATCCGCAAGGGGTGATTA 68
QY 186 CATTTTGGCTATCCCGCGGGGCAATTTTGGCCATCTATGATGAATCTGACGCTTGA 245
Db 69 TATTTTGGTATCTCGTGGGGCAATTTACCAATTTATGACGACTGTATAGGTGA 128
QY 246 AGCGGCGGGGAAATGAGCATATTTTGGTGGCCATGAACAAGAGGCTTCCAGCGGC 305
Db 129 AGCAATCTGTGCTATTAGCATATTTCTGTAGACACGAAACAAGGCGCCATGCGC 188
QY 306 GATGGGTAATGCCAGAGCCACAGTAAAGTGGAGTTTGTTCGGTACATCTGACGAG 365
Db 189 AGACGGTACGCGCGCCACTGGGAGGAGTGAAGTATGCTTTGGCACTTGGCCAGG 248
QY 366 GGGACTAATCTGTGACCGGCAATGGCCATGGCCATTTGGACTGGTCCCATGTGTGT 425
Db 249 AGCACTAATCTGTGACAGGAGTGGCCACCGCTACATGATGATCAATCCCATATGT 308
QY 426 GATTAATGAGAGTGGGCGGTGCGCATGATGTGACGATGCTTCCAGAAATTGACAT 485
Db 309 GGTACGGGGCAGGTAGACGTCCTGCTCATTTGTACAGATGCTTTCAGAAACGATAT 368
QY 486 TTTTGGCATCACTTAACGATGTTTACGATCTCTATGTGTGATGAGTGGCGGAGAT 545
Db 369 TTACGGCATTAAGCAACCATAGTAAAGCATCTTATGATGAGCTGACCCCTAAAGACAT 428
QY 546 GCGTCGATTTGTTACTGAGGCTTTTCATCTTGTGACACCGGTGCTCCGCGCGGTTT 605
Db 429 GCGCGGAATTTGCGCGAAGCTTTTCACTGCTAGCATGGGCGCAACAGGCGCATTTCT 488
QY 606 GATCATATTTCCCAAGATGTGGGCTTAAAGAAATGTGATCATTTCCCTGACCCCGG 665
Db 489 GATGATGTTCCTCAAGATGTGATGCTTTAGAAATTTGATGATGATGATGATGATGAT 548
QY 666 TACGCTAATCTAACGGGGTTATCGGCCCAAGCTTAAAGTAAATCCCGCAAAATTAAGC 725
Db 549 TTCAATGATGATTAAGGGGTTATCGGCCCAAGCTTAAAGGAAATCCCGCAAAATTAAGC 608
QY 726 GGCATATGCAATTTGTGAGACAGCGCAAAATCTTGTGCTTACGATGAGGAGGCGAT 785
Db 609 TGCATATGATGATTAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 668
QY 786 GCGCCGCAATGGCCATGCGCCAGGTGAGAAATTTTCCGAAAGGTTCCAGTTGCGGATAC 845
Db 669 GCGACCTGCTGCTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 728
QY 846 AACCAACCTGATGAGAAATTTGGGCTTTTGAAGAAATCAATCCCTTGGGAGGATGAT 905
Db 729 CACCACTTAATGAGGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 788
QY 906 GGGTATGATGACGCAACCGCTATGCAACTTTGCGCTGACGAGATGATGATGATGATG 965
Db 789 GGGAGATGACGCGCAAGGCTTACGTAATTTGCTATGATGATGATGATGATGATGATG 848
QY 966 AGTGGGGGCGCGTTTGAAGCAAGGCTTAACTGAGCAAAATTAAGCAATTTGCTAGCGCGC 1025
Db 849 CCGTGGCGCTGATTTGAGAGATGCGCTTACAGGCAAGTAAAGCAATTTGCGCTCCCGGC 908
QY 1026 CAAATTAATTCATTTGATGATGACCGCGGAGGTGGAAAAAACAAGGCTCCCGCATGT 1085
Db 909 CAAATTAATTCATTTGATGATGATGACCGCGGAGGTGGAAAAAACAAGGCTCCCGCATGT 968
QY 1086 GCGCATTTGGGGGATGTAACCGCATGTTTGAAGAAGCTTTTGAAGCGGGCGCGGAGAT 1145
Db 969 GCGCATTTGGGGGATGTAACCGCATGTTTGAAGAAGCTTTTGAAGCGGGCGCGGAGAT 1028
QY 1146 GGATTAACCCCAATCCCATACCAAGGCGATGTTAAATGATGATGATGATGATGATG 1205
Db 1029 AGGCGTTAAGGCTACCCCTAATCAAAACCAAGAAATGTTAATCTTGTAAACGTTGGCG 1088
```

```
QY 1206 GACCGATTACCCCTCCAGGAGGCCCACTATGAGGATTAATTTGCCCCCGGAGAGGAT 1265
Db 1089 CGATGATATCTTAAATATGATGACAGCACCCGACGATTTCAACCAAGAGTAT 1148
QY 1266 ACAGAAATTTGATGCGAGGCGCCCGATGCTTACTACCAACGATGAGGAGAACCA 1325
Db 1149 TGTAAATGATATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1208
QY 1326 AATGAGGCGGCGGAGTTTGAACATATGCCCCCGGATGATTTTCAATGCTGCTT 1385
Db 1209 AATGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1268
QY 1386 GGGTACGATGAGGCTTTGTTTACCTGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444
Db 1269 GGGAGGATGAGGCTTTGCTTACCTGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1328
QY 1445 AGCGTATTTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1504
Db 1329 AGAGTATTTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1388
QY 1505 CCGAGCCAGATGACATCCAGGTTAAATCTATTTCTCAATTAAGGTTGGAGAGGAGAT 1564
Db 1389 ACTTGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
QY 1565 GGTGCGTACGTGCGCAAACTTTCTACGAAAGGATTTCTGCTTCAACATGTCCTCA 1624
Db 1449 GGTGCGCGCAGTGGCAACAGCTTTCTATGATGAGGATTTCTGCTTCAACATGAGAT 1508
QY 1625 GGGATGCGCAGATTAATATCTCTGTGAAGGCTTATGAGGATGATGAGGATTTACTGCG 1684
Db 1509 AGGATGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
QY 1685 CAAGCGGAGATTTTGGCGCGCGGATGCGCGGATGCGCGGATGCGCGGATGCGCGGAT 1744
Db 1569 CGATGCGATCAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
QY 1745 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1804
Db 1629 CTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
QY 1805 TAAATGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1864
Db 1689 CAAGCTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1745
QY 1865 GAGTGCACCATGTCGCAAAACCAAAATTTATCATACCATGCTTCTGTTGTTGTTG 1924
Db 1746 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
QY 1925 AGCCAACT 1933
Db 1806 GACGAGCT 1814
```

```
RESULT 3
US-10-369-493-42620
; Sequence 42620, Application US/10369493
; Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
PRIOR FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360,039
NUMBER OF SEQ. ID NOS: 47374
; SEQ ID NO 42620
```

; LENGTH: 1815
 ; TYPE: DNA
 ; ORGANISM: Anabaena PCC7120
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1815)
 ; OTHER INFORMATION: unsure at all n locations
 US-10-369-493-42620

Query Match 43.0%; Score 841.8; DB 16; Length 1815;
 Best Local Similarity 67.3%; Pred. No. 8,66-272;
 Matches 1215; Conservative 0; Mismatches 586; Indels 4; Gaps 2;

QY 127 CAACGGCAACCCGGGCTTATATCTGATGATGCTGAAACGCCATGGGCTCAACAC 186
 DB 10 CAACGGGTACTGCTGCTTGCCTACTTATGATGCTGCTGCCCAAGCGCTTGAGTAT 69
 QY 187 ATTTTGGCTATCCCGGGGGGCAATTTTGGCCATCTATGATGATGATGATGATGATGAT 246
 DB 70 ATCTTTGGTATCTCTGCTGGGGGCAATTTCTGCGATTTATGATGATGATGATGATGAT 129
 QY 247 GCGGCGGGGGAATTTGACATATTTTGTGCGCCATGAAACAAGAGCTTCCCATGCGCG 306
 DB 130 GCAACTGGCAGCTTAAACATATTTCTAGTGAACACGAAACAAGAGAGCTCATATGCTCT 189
 QY 307 GATGCTATGCGCAGAGCCACAGTAAAGTGGAGTTTGTTCGATGATGATGATGATGATGAT 366
 DB 190 GATGCTATGCTGCGCAACCGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 249
 QY 367 GCGACTTACTGTCAGCCGATTCGCAATGCCATGCCATTTGGACTGCTGCTGCTGCTGCTG 426
 DB 250 GCAACTTACTGCTGTCAGAGTATGCTACAGCTTCAATGATGATGATGATGATGATGATGAT 309
 QY 427 ATTATGAGAGAGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
 DB 310 GTTACTGGGCAAGTACCCCGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
 QY 487 TTGCGATCACCTTACCCATGCTTAAAGCACTCTATGCTGCTGCTGCTGCTGCTGCTGCTG 546
 DB 370 TACGCTATTTACTCTACCATTTGTAAGCACTCTATGCTGCTGCTGCTGCTGCTGCTGCTG 429
 QY 547 GCTGCTATTTACTGAGGCTTTTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
 DB 430 GCGCAATTTGTCGCAAGGATTCACATGACGACACAGGCGCTGCAAGGCGCGAGTGTG 489
 QY 607 ATCGATTTTCCCAAGAGATGGGCTTGAAGATGATGATGATGATGATGATGATGATGATGAT 666
 DB 480 ATTGATGCGGAAAGATGCTGCTTTTGAACAATTTGATGATGATGATGATGATGATGATGAT 549
 QY 667 GACGTTAATCTACCGGGTTATCGCCCAAGGTTAAAGTAAATCCCGCAACAATTAATGCG 726
 DB 550 TCAATTAATTAACGGGATATCGCCCAAGGTTAAAGGTTAAATCCCGCAACAATTAATGCG 609
 QY 727 GCATTGCAATTTGTCAGACAGGCGCAAAATCCCTTGTCTACGTAAGGGGGAAGGGGCGATC 786
 DB 610 GCGATTGAATGATCGGAAAGCGCTGCTTATCAATGCTGCTGCTGCTGCTGCTGCTGCTG 669
 QY 787 GCGCCCAATGCCCATGCGCAGGTGCAAGAAATTTGCGAAAGTTTCCAGTTTCCCGGTAAACA 846
 DB 670 GCATGCGGCTCATGCAAGAAATCAAGAACTCGCGAGTTAATTAATCCCGTCACT 729
 QY 847 ACCACCTCATGAGGATTTGGGGCTTTTACGAAACCATCCCTTGTGCTGCTGCTGCTGCTG 906
 DB 730 ACCACCTCAATGGGATCGGGGCTTTTATGAACCAATCTTATCTTGGGGATGTTG 789
 QY 907 GGTATGCAATGGCCACCGCTATGCAACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
 DB 790 GGGATGCAAGCAGCTTACCTTACCTTGTGATGCTGATGATGATGATGATGATGATGATGAT 849
 QY 967 GTGGGGGCGCTTTGACGACCGGGTTAATGCGAAACTGACGATTTTGTCTAGCCGCGCC 1026
 DB 850 GTGGGTGCTAGTGTGATGACCGTGTACTGCGCAANTTGAATGATGATGATGATGATGATG 909

QY 1027 AAAGTAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
 DB 910 AAAGTAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
 QY 1087 CCCATTGTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146
 DB 970 CCCATTGTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1029
 QY 1147 GATTAACCCCAATCCCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
 DB 1030 AGTGCCTAAATCTACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1089
 QY 1207 ACCGATTAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1266
 DB 1090 GAGGATTAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCA 1149
 QY 1267 CACGAATTTGCTGCGCAGGCGCCGATGCTTACTACACCATGATGATGATGATGATGATGAT 1326
 DB 1150 GTGGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
 QY 1327 ATGTTGGGCGGCGGCTTTTGAACATGAGGCGCGCGATGATGATGATGATGATGATGATGAT 1386
 DB 1210 ATGTTGGGCGGCGGCTTTTGAACATGAGGCGCGCGATGATGATGATGATGATGATGATGAT 1269
 QY 1387 GGTACGATGAGGCTTTGCTTACCTGCGCATGAGGCGCAAGTGGAGTGGGAGCGAG 1446
 DB 1270 GGAATGATGAGGCTTTGCTTACCTGCGCATGAGGCGCAAGTGGAGTGGGAGCGAG 1329
 QY 1447 C-GGTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
 DB 1330 CAAGTAACTGTTAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
 QY 1506 CTAGCCAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
 DB 1390 GCTGCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
 QY 1566 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1625
 DB 1450 GTAGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1509
 QY 1626 GCGATGCGCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1685
 DB 1510 GGGATGCGCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1569
 QY 1686 AAGCGGGAATTTGGCGCGCGATGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1745
 DB 1570 CACAGAGACCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629
 QY 1746 ATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805
 DB 1630 TTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689
 QY 1806 AATGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1865
 DB 1690 AATGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1746
 QY 1866 GAGTGCACCAATTTGCCCAACCCCAATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1925
 DB 1747 TATTTGACGATTTGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
 QY 1926 GCCAA 1930
 DB 1807 ACTTA 1811

RESULT 4
 US-10-282-122A-41727
 ; Sequence 41727, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl

/ APPLICANT: Haselbeck, Robert
 / APPLICANT: Ohlsen, Kari
 / APPLICANT: Zyskind, Judith
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John
 / APPLICANT: Carr, Grant
 / APPLICANT: Yamamoto, Robert
 / APPLICANT: Foreyth, R.
 / APPLICANT: Xu, H.
 / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 / FILE REFERENCE: ELITRA 0344
 / CURRENT APPLICATION NUMBER: US/10/282,122A
 / CURRENT FILING DATE: 2003-02-20
 / PRIOR APPLICATION NUMBER: 60/191,078
 / PRIOR FILING DATE: 2000-03-21
 / PRIOR APPLICATION NUMBER: 60/206,848
 / PRIOR FILING DATE: 2000-05-23
 / PRIOR APPLICATION NUMBER: 60/207,727
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: 60/230,335
 / PRIOR FILING DATE: 2000-09-06
 / PRIOR APPLICATION NUMBER: 60/230,347
 / PRIOR FILING DATE: 2000-09-09
 / PRIOR APPLICATION NUMBER: 60/242,578
 / PRIOR FILING DATE: 2000-10-23
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-11-27
 / PRIOR APPLICATION NUMBER: 60/257,931
 / PRIOR FILING DATE: 2000-12-22
 / PRIOR APPLICATION NUMBER: 60/267,636
 / PRIOR FILING DATE: 2001-02-09
 / PRIOR APPLICATION NUMBER: 60/269,308
 / PRIOR FILING DATE: 2001-02-16
 / Remaining Prior Application data removed - See file wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 78614
 / SOFTWARE: Patent in version 3.1
 / SEQ ID NO 41727
 / LENGTH: 1728
 / TYPE: DNA
 / ORGANISM: Versinia pectis
 / US-10-282-122A-41727

Query Match 21.2%; Score 415; DB 13; Length 1728;
 Best Local Similarity 55.8%; Pred. No. 4,7e-128;
 Matches 929; Conservative 0; Mismatches 705; Indels 31; Gaps 6;

QY	172	CATGGGGTCAAA	CATT	TTTGGCTATCCCGCGGGGCAATT	TGGCCATCTATGATGAA	231
DB	49	CAGGCGGTAA	AGCATGTAATTCGTTATCCCGCGGGGCGTACT	GTATATTACGATGCC	108	
QY	232	CTGTACCGCTTGA	AGCGGCGGGGAAATTGAGCATATTTTGGTCCGCATGAA	CAAGGA	291	
DB	109	CTGCAC-----	ACGGTCGAGGACATCGATCACTGCTGGTGGCCATGAA	CAAGGT	159	
QY	232	GCTTCCATGCGGCGGATGGATG	CCAGAGCCACAGGTAAAGTGGAGTTT	TTTGGT	351	
DB	160	GCGGTACATGCGCGATGGTATG	CGGGGCAAGCGGAGGTGGTGTGACTGGTC	219		
QY	332	ACATCTGACACAGGGGCGCAT	TAATTTGGTACCCGCATTTGGCCATGGCCATTTGAC	CTGC	411	
DB	220	ACGTCTGCTCTGGTGCACACAT	GCATACCGGTATGGACCGCTATATGATTA	ATCA	279	
QY	412	GTGCCATGTGTGATTA	CTGAGAGGTGGCCGTGCCATGATTTGGTAGCATGCTTTC	471		
DB	280	GTGCAATGTGTGCTTTCAGG	CCAGTGCAGATGATTTGATGACCTTC	339		
QY	472	CAGGAAATTTGACATTTTGG	ATCATCTTACCGATCTTAAAGACACTCTCTATGTGTACGT	531		
DB	340	CAGGATGTGACATGTGGGAT	CTCCCGCCAGTGTCAACATAGCTTCTTGGTTAAG	399		
QY	532	AGTGGCGGATATGGCT	CGCATTTGTACTGAGGCTTTCATCTTGTAGACCGGATGT	591		
DB	400	CGCAAGAAACATCCGAT	GGTGTGAAAAAGCTTTTATCTGGATCCATCGGCGCT	459		

QY	592	CCCGGCGGGTTT	GTATATCCCAAGATGTGGGCTTAGAAGATGTGATCAT	651
DB	460	CCCGTCCGGTGTAT	ATGACTTCCCAAAAGCATTTTGGCCCGAGATGAAATGCT	519
QY	652	CCCTTCAGCCCGGT	ACGTAACTACCGGGTATCCGCCACGTTAAAGTAAATCC	711
DB	520	TACGCTATCCAGAG	CAAGTCAAGTTTCCCTATATCTACGGTGAAGGCAAGCT	579
QY	712	CGAAATTTATGGG	ATTTGCAATTTTGTAGACAGGCCGAATCCCTTGTCTACGA	771
DB	580	GGGCAATTAAGCG	GTGCTGCAACGATTTATGGCTTAAACGGCCATATATGATGC	639
QY	772	GGGGAGGGGCGAT	TCGCGCAATGCGCATGCGCAGGTGAGAAATTTGGGAAAGTTC	831
DB	640	GGTGTGTGCGAT	TAACGCGCTTGCAATGAAGATTAATTTGGAGAAAGCTG	639
QY	832	CAGTTCGGGTAA	CAACACCTGATGGAAATGGGGCTTTGACGAAACATCCCTT	891
DB	700	AACCTACCGGTAA	CCAGCTCTGATGGGCTAGGCAAGTTCCCTGTACGCAATCCCA	759
QY	892	TCGGTGGGTAT	GTGGATGACAGGCAACGCTATGCCAATTTGGCGGTACGAGATG	951
DB	760	AGTGTGCGCAT	GTGGGATGCACTGTGTAAGCCAAATTTGGCATGCAATACT	819
QY	952	GATTTGTGATG	CAATGGGGGCGGCTTTCAGACACCGGTTAACTGGCAAACTAGAGAA	1011
DB	820	GATCTGATTTT	TCGCGGTGTCGCTGTCGCTTTTGAAGATGATCCCAATATTTAGCAAA	879
QY	1012	TTTGTACCGCG	CCAAAGATTAATTCATTTGATGACCCCGGAGAGTGGGAAATAAC	1071
DB	880	TATTCCTTATG	CCACAGTCTACATTTGATGATGATGATGATGATGATGATGATGATG	939
QY	1072	AGGGCTCCGAT	GTGCGCATTTGGGGGATGATGCGCATTTTGAACAGCTTTTGCAG	1131
DB	940	GTCACCGCGAT	ATCCGATTTGGGGATGCAAAACAGTACTGACAAATGTTAGAT	999
QY	1132	CGGGCCCGGAA	TGTGATTAACCCCAATCCCATCCCATCCAGGATGTTAAATCGC	1191
DB	1000	TTGTGTGCGAG	GTGCGCTGTACAAAGATTTGATGATGATGATGATGATGATGATG	1059
QY	1192	ATTGATCAT	TGGGAGCCGATTAACCCCTCCAGGTCCCGCATATGATGATGATG	1251
DB	1060	ATTGACAAAT	GGGCGACCCGTAATTTGCTGGGTACAAACAGAGTGGGAAATCAAG	1119
QY	1252	CCCCAGAGGT	ATGACAAATTTGGTCCAGGCCCC--CGATGCTTACTACACACC	1308
DB	1120	CCACAAGCGGT	ATGATGACATGCTTGAATTAAGGCGATGCTGATACCTTACTCT	1179
QY	1309	GATGTGGCA	CAACACCAATTTGGGGGCCCC--ATTGTTGAACATGCCCCCGCCA	1365
DB	1180	GATGTGGCCAA	ATATGATTTGGGCTCTTTACTATTCCTTTGATTAAGCAGCTCGT	1239
QY	1366	TGATTTTCA	AGTCCGCTTGGGTACATGAGGCTTTGTTTACTGCTCCGCAATGGAGGC	1425
DB	1240	TGATTTAACT	CCGCTGGCTTGGGCAAGATGGGCTTTGGCTTGCAGCGGCTTGGGGT	1299
QY	1426	AAAGTGGAGT	GGGAGAG--AGCGTCAATTTGATCATGTGAGATGCCAGCTTCCAAAG	1484
DB	1300	AAATTTGCTT	GTATATGAACCGTGTGTGTGATGACGGGTATGGCATATCCAGATG	1359
QY	1485	AATCTTCAG	AACTGGAAACCTTAGCCAGTACGACATCCAGGTAAATCTATTTATCTC	1544
DB	1360	AATATTCAG	AGATTAACAACCGCATCTGCAATTAATTTCCAGTGTGGATTAACCTG	1419
QY	1545	AATTAACG	TTTGGAGGAGATGTGCTCACTGACAACTTTCTTACAAAGACGTAT	1604
DB	1420	AATTAACCG	GTATATGGGATGTGGAACAATGGCAGACATGATTAATTCAGCGCCGCAAT	1479
QY	1605	TCGTCTTCA	CACTGTCCAGGCGATGCGACATTAATCTCCTGTGAGAGCTATGCG	1664
DB	1480	TCGATGCTG	ATATG--GATTCGCTGCTGATTTTGTCAAGTGGCTAAAGCTTATGCG	1536

QY 1306 ACCGATGTTGGAGCAACACCAATGTTGGGCGCCAGT---TTTGACATATGGCCCCGC 1362
 Db 1294 GGGGTTGATAGTCAACATCAGATGTGGGGGCGAGTTTCATCTTACAGAAAGCCGCT 1353
 QY 1363 CGATGATTTCCAGTGTGGCTTGGGATGATGGGCTTTGTTTACCTGCGCCATGGGA 1422
 Db 1354 ACTTGGCTCACTCAAGCGGTGTGGGACCAATGGGTTTGTATTTCCGCGCCATGGGC 1413
 QY 1423 GCCAA-AGTGGAGTGGGGGAGCAGCGGCTATTGCAATCAGTGGAGATGCCAGCTCCAA 1481
 Db 1414 GCCAAGATTGCTCCCGCGAGGCAAGTCTGGGCGATGACGCGGATGGCTGTCTCAG 1473
 QY 1482 ATGAATCTTCAGAACTGGGAACCTAGCCAGTACGACATCCAGTTAAACTATTAAT 1541
 Db 1474 ATGACCAACAGAGTGTGGCTAGTGCCTATCGAGGGGCGCCGATCAAGTGGCGTGT 1533
 QY 1542 CTCATTAAGGTTGGCAGGGATGTGTGCTCAGTGGCAACAACCTTTCTACAGAAAGCT 1601
 Db 1534 ATCAACAAGCGCACTGGGATGTGGGCAATGCGCAAGCTTGTATTCAGAGAGCT 1593
 QY 1602 TATCTGCTTCTAAGATG-----TCCAGGGCATGCGCAGACATTAATCTCTGT 1652
 Db 1594 TACTGGCAGACCGACTGGCCACCACTCACACCGCATTCGGAATTTGTGAAGTTGCC 1653
 QY 1653 GAAGCTATGGCATGAAGGATTAATCT---GTGGCAAGCGGGAAGATTGGCCCCGCG 1709
 Db 1654 GAGGGGTGGGTTGGCTGGATTTGGCTTGGAGTGAAGAGAGTGTGATGTATC 1713
 QY 1710 ATCCGCAATGCTAGCCCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1769
 Db 1714 AATCAGGCGGAGGCGCATCAACAACCGCTGTGGTGAATGACTTATGTGTGGCGCGAT 1773
 QY 1770 GAAAATCTGTACCTATGATTTGCCCGCATGAGTAAATGCCCAATGCTAGT 1823
 Db 1774 GCGCAGTGTGGCCGATGATGTCGCTGGGGCCAGCAAGAGATCCAAAGCT 1827

RESULT 6
 US-10-282-25560
 ; Sequence 25560, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haseibeck, Robert
 ; APPLICANT: Ohlsen, Karl
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Truick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27

QY PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 25560
 ; LENGTH: 1866
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium avium
 US-10-282-122A-25560

Query Match 19.4%; Score 379.8; DB 13; Length 1866;
 Best Local Similarity 54.3%; Pred. No. 3.5e-116;
 Matches 938; Conservative 0; Mismatches 757; Indels 34; Gaps 7;

QY 109 CTGAAATCTGTTCTAGCCCAACGCAACCGGGCTTATATCTGATGATAGCTGAAG 168
 Db 100 CCGAAAGCATTTGGGCGCCGCAAGTTACTGTGCAACAATCGGTATCCGATCGCTGAG 159
 QY 169 CGCCATGGGTCAACAACATTTTGGCTATCCGGGGGGGCAATTTTGGCCATCTATGAT 228
 Db 160 GAGCTGGCGTGAAGTATCTTCGCAATTCGCGGGCGCGGTGTGCTGCGCTATGAC 219
 QY 229 GAACTGATCCGCTTTGAAGCGGGGAAATTGACATATTTGGTGGCCATGAACA 288
 Db 220 CCGGTGTGACT-----CCAGAGCTCGGCAAGTGTGATCGGCAAGGCTGTCTGAC 270
 QY 289 GAGCTCTCCATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 348
 Db 271 GATGCGGTGATCCCGCGCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 330
 QY 349 GATACATCTGACCAAGGGGCACTAATGATGATGATGATGATGATGATGATGATGATGAT 408
 Db 331 GCACTCTCGGTCCGGCGGCAACCACTGATGATGATGATGATGATGATGATGATGATGAT 350
 QY 409 TCGGTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 468
 Db 391 TCGATCCCGCTCGTCCCGCTCAACGCGAGTGGAGCAAGCTGATCGGCAACGAGCGC 450
 QY 469 TTCCAGAAATTTGATGATTTTGGCATCAGCTTACGATGATGATGATGATGATGATGAT 528
 Db 451 TTCCAGAGGCGCAACATCTCCGATCAGATGATGATGATGATGATGATGATGATGATGAT 510
 QY 529 CGTATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
 Db 511 CGTCCGCGCAAGAAATCCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 570
 QY 589 CGTCCGCGCGGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648
 Db 571 CGCCCGCGCGGTGTCTGCAATCCCAAGAGATGATGATGATGATGATGATGATGATGAT 630
 QY 649 ATTCCCTGCAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
 Db 631 AG-----CTGGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 684
 QY 709 CCGGCAAAATTAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768
 Db 685 AGCCGCGAGATCCGCAAGCGCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 744
 QY 769 GTAGGGGAGGGGATGCGCGGCAATGCGCAATGCGCAATGCGCAATGCGCAATGCGCAAT 828
 Db 745 GTAGGGGCGGTTCATCCGCGGAGGCGCAAGGATGATGATGATGATGATGATGATGATGAT 804
 QY 829 TTCCAGTTGCGGTTAAACAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 888
 Db 805 ACCGCAATCCCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 864
 QY 889 CTTTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948

Db 865 CAGCACTGGGCGATGCCGGGATGCAACGACCCGTCCGGCGGTGGCGGCCCTTGCAAGCC 924
 Qy 949 TGTGATTTGATTTGATGCAAGTGGGGGCGCCGTTTTCAGCAACCGGGTACTGGCAACTAGAC 1008
 Db 925 AGCGACTGTGATGCGGCTGGGTACGGGTTTCATGACCGGGGTGACCGGCAAGCTGAC 984
 Qy 1009 GAATTTGTAGCCGGGCCAAAGTAAATTCATTCATTCACATCCCGGGGGAGGTGGGAAA 1068
 Db 985 ACCTTGGCCCGGAAAGCCAGATCATCCACGCCACATCCACCGCGCGAGATCGGCAAG 1044
 Qy 1069 AACAGGCTCCCGATGTTGCCATTTGTGGGGATGTAGCCCATGTTTGAACAGCTTTTG 1128
 Db 1045 AACGGGATGCGGAGCGTCCGATGCTGGCGACGTCACAGCGCTCATCCGCGAATGTC 1104
 Qy 1129 CAGCGGCGCGGGAAATTTGATTTACCCACCATTCCTCCATACACCCGAGATGTTAAT 1188
 Db 1105 GAGATGCTCGCCACGACGAGGCGCGCCGCAATCTCCACATCCGCGCATGCTGGCCCTAC 1164
 Qy 1189 CGCATTTGATTCATTTGGCGGACCGATTAACCCCTCCAGGTCGCCCTACTAGAGTACTATT 1248
 Db 1165 CTCACGACGCTCCAGTCCACTATCCGCTGAGCTACGGGCGCGACGACGAGGCTC 1224
 Qy 1249 GCCCCCGAGAGTATGACGAAATTTGTCGCAAGCC--CCGATGCTTACTACACC 1305
 Db 1225 GGGCCCGAGTACGTCATGAGAAAGCTGGCGCAGATCCCGGCGCGACGCTGTACGTG 1284
 Qy 1306 ACCGATGTTGGGACCAACCAATGTTGGCGGCGCCAGT---TTTGAACAAATGGCCCGC 1362
 Db 1285 GCGGCGTGGGTACGATCAATGTGGCGGCGCCAAATTCATCTCTTCGAGAAAGCCGCG 1344
 Qy 1363 CGATGATTTTCAGTGTGCTGGGTGAGTACGATGGGCTTTGGTTTACTCTGCGCCATGGA 1422
 Db 1345 ACTTGGCTCAATTCGCGGCGCTGGGACCAATGGGGTTTGGCATTTCCGCGCGCATGGGC 1404
 Qy 1423 GCCAAATGGG--AGTGGGGGACGACGCGTCAATTTGCATTCATGAGATGTCAGTTCCAA 1481
 Db 1405 GCCAAGATGGCGGTCCGAGGCGCGAGGTGTGGCGATTCAGCGCGCGGTGTTTCCAG 1464
 Qy 1482 ATGATCTTTCAGAACTGGGAAACCTAGCCCAAGTACGACATTCAGTTAAACTATATT 1541
 Db 1465 ATGACCAACGAGAGCTGGCCACCGCGCGCTGAGAGGCAATCCGATCAAGGTGGCGCTG 1524
 Qy 1542 CTCATTAACGGTTGGCAGGGGATGTGTGTAGTGGCAACAAATTTCTTCAGAAAGCT 1601
 Db 1525 ATCAACAAACGCAACCTGGGATGTGGCCAGTGGGAGAACCTGTCTTCAGAAAGCCG 1584
 Qy 1602 TATTC-----TGCTTCAATGATGCCAGGGGCATGCGACATTAATCTCTCTGT 1652
 Db 1585 TATTCGCAACGATCTGGCCACCCACTCCGACCGCATCCCGGACTTCGTCAGCTGCC 1644
 Qy 1653 GAAGCTATGGCATCAAGGGTAT--TACTGTGGCAAGCGGGAAATTTGGCCCGCGC 1709
 Db 1645 GAGGCGTGGGCTGCTGTCGATGCGTTGCGAGCGTGAAGAGCGTCGTCAGCTGATC 1704
 Qy 1710 ATTCGCGAAATGCTAGCCCAATGTCCTGTGTGATGATGTGTGTCACAAAAGAT 1769
 Db 1705 AACGCTGCCCGGCGGATTAACGACCGCGCGGTGTGATTCACCTTCATGTCGGCGGAC 1764
 Qy 1770 GAAAACTGTACCTATGATTTGCCCGGCGCATGAGTATGCGCAATGC 1818
 Db 1765 GCGCAGGTGTGGCGATGTGGCGCGGACGAGCAACGAGATCC 1813

RESULT 7

US-10-282-122A-41305
 ; Sequence 41305, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELTRA, 034A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,846
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,576
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 41305
 ; LENGTH: 1722
 ; TYPE: DNA
 ; ORGANISM: Vibrio cholerae
 ; US-10-282-122A-41305
 Query Match 19.4%; Score 379.6; DB 13; Length 1722;
 Best Local Similarity 53.9%; Pred. No. 3,9e-116;
 Matches 917; Conservative 0; Mismatches 764; Indels 19; Gaps 6;

Qy 128 AACGCCAAACCGGGCTTATATCTGATGATAGCTTGAAGCCGATGGGCTCAACACA 187
 Db 5 AATGCTATCTGGCCGACAGATGATCGTCAATCTGATCAATGAAGTGTGGACAAA 64
 Qy 188 TTTTGGGCTATCCGGCGGGGCAATTTGGCCATCTATGATGAACCTGACGTTTGAAG 247
 Db 65 TCTTGATATCCGGCGGCTTACGATCTGACATCTACGATGCCCTTAC-----GAAA 118
 Qy 248 CGCGCGGGGAAATTGAGCATATTTGTGCGCCATGAACAAAGAGCTTCCGATGGCGCG 307
 Db 119 AAACGACCAATTAACACGTTCTGTTGTCAGAAACGCGCAACACATATAGCTG 178
 Qy 308 ATGGGATATGCCAGACCCACAGTAAAGTGGAGTTTGTGGTACATCTTGGACCAAGGG 367
 Db 179 ATGGGTATGGCGGCCCACTGGAAAAACGGGAGTAGTGTCTGTCTGCAAGTCCGGCG 238
 Qy 368 CGACTAATCTGTGACCGGCAATGGCCATGGCCATTGGAATCGGTGGCCATGGTGTGA 427
 Db 239 CCACCAATACGTTGATGAGATTTGAACCGGCTATATGATTCATTCATTTGGA 298
 Qy 428 TTAATGAGAGGTGGCGCGTGCATGATTTGTAGCGATGCTTTTCAAGAAAATTGACATT 487
 Db 299 TCTCCGATACGTTACCACTATCTTATTTGATGAGATTTTCMAAGTGCATATCG 358
 Qy 488 TTGGCATACCTTACGATCGTTAAGCACTCTATGTGTGATGATGATGCGCGGATATG 547
 Db 359 TCGGCGTGTCTGTCCATCTGTTAAACACAGCTTTTGTGAAAAAACCGAAGTATTC 418
 Qy 548 CTGCAATTTATCTGAGGCTTTTCATCTTGTACGACCGGTGTCGCGCGCGGTTTGA 607

Db	419	CTGAACCATCAAAAAAGCGTTTATATATGCGCTCAACGGGGCGTCTGTCTTGTAAGTA	478
OY	608	TCGATATTTCCAMAGATGTGGGCTTATGAAGAAATGTAGTACATTCCTCCGACCCCGGTG	667
Db	479	TTGACGTGCAAAAAGATGTGATGAGCCCGCTTAATAAGCTGCTTATGAGTATCCAGAAA	538
OY	668	ACGTTAATCTACCGGGTATCGCCCCACGTTAAAGTATATCCCGACAAATTAATGCG	727
Db	539	CCATTAAATGCGCTTTTACAAACCAACACGCGTGGCCACAAAGACAGATCAAAAAG	598
OY	728	CATTGCAATTTGTTGAGCAGGCCAGAAATCCCTTGCTCTACGTAGGGGGAGGGCGATCG	787
Db	599	GGTTACGGGCACTGCTCCATGCGAAAAAGCCAGTACTGTACGTCGGGTGGCGGGCGATT	658
OY	788	CCGCCAATGCCATGCCACAGTGCAGGAATTTGGGAAAGTTCCAGTTGCCGTTACAA	847
Db	659	TTGCTAAATGCAGATAGCAAAATTCGCCAATTCGACAGAGCTTTAATCTGCCCGTTTA	718
OY	848	CCACCTGATGGAAATTTGGGGCTTTTGAAGAAACATCCCTTGCGGGGTATGTTGG	907
Db	719	GTACTTTGATGGATTGGGCGTTTTTCTCGGACCTCAATAAACTCATGCGGTATGCTTG	778
OY	908	GTAATCATGCGCACCGCTATGCGCAATTTGCCGTACGGCAATGTATTTGTTGATTGCA	967
Db	779	GCATCATAGTGTGTATGAGCCAAATATGGCGATGCACAAATGCATGCATTTTGGGG	838
OY	968	TGGGGGCGCGTTTGACGACACGGGGTACTGGAAACTGAGCAATTTGCTAGCCCGCA	1027
Db	839	TAGGTGTGCGCTTTGATGACGTACGACTTAACAACCTGAGAGTACTGCCCCAACGGA	898
OY	1028	AAGTAATTCATTGACATTCGACCCCGGGAGGTGGGAAAAAACAAGGCTCCCGATGTGC	1087
Db	899	AAATCATGCATTCGATATTTGACCCGTATGCAATTTCTAATAAACGTGAAGTCGATCTGC	958
OY	1088	CCATTGTGGGGATGTAGCCCATGTTTATGAACAGCTTTTGCAGCGGGCCGGGAATTGG	1147
Db	959	CGATTGTGTGTTCTGCCGACCAAGATTTGGAGGCGATCCTCAAGCTTCTGAAAGAGAGC	1018
OY	1148	ATTACCCACCCATCCCATACCAACCCAGGCACTGGTTAAATGCAATGATCATAGGCGGA	1207
Db	1019	CTGAGCGCAATGATGCGGACGATTTGGATAGTGTGTAATGATCCAACTGTGGCGTG	1078
OY	1208	CCGATTACCCCTCCAGGTGCCCACTATGAGATACTATTGCCCCCAGAA--GGTAA	1267
Db	1079	ATCGCAATGTTTGGCTTATGAACCTCAGACGACAGGATTAAGCTCAGCAAGATTTG	1138
OY	1265	TACACGAATTTGATGCGCAGGCGCCCGGAGTCCCTACTACACACCGATGAGGACAAAC	1327
Db	1139	AAAGCGCTTATTAATAACGGAAGGTAAAGCAATTTGGCTCGGATGTGGGCGACCC	1199
OY	1325	AAATGTGGGCGGCGC--CCAGTTTGTGAACAATGGCCCCCGCGCATGATTTCCAGTGTG	1381
Db	1199	AAATGTTTGGCGCACTTACTATTCGTTTGAATTAAGCCCGTCAATGATTAACCTGTGGCG	1255
OY	1382	GCTTGGGTACGATGGGCTTTGTTTAACTGCGCGCATGGAGCCAAAGTGGAGTGGGGG	1444
Db	1259	GGCTAAGGCAACATGGGTTTGGTTTAACTGCGGGGATGGGCGTGAAATTCGCAAGCCAG	1318
OY	1442	ACG-AGCGGTATTGTCATTCAGTGGAGATGCGAGTTCCAAATGAATCTTCAGGAATCG	1500
Db	1319	AAGAAGAGTACTGTCGTGACCGGAGATGCGAGTATCCAGTGAATTTCAAGAGCTAT	1378
OY	1501	GAACTTAGCCCATGACGATCCAGGTTAAAATCTATTAATCTCAATAACGTTGGCAGG	1560
Db	1379	CGACCGCACTGCAGTACGACATTCACGTAATAATATTAACTCAACACCGTTCTTGG	1438
OY	1561	GGATGTCGTGACGTGGGACAACTTTCTACGAAGAAGTATCTGTCTTCAACATG	1620
Db	1439	GAATGTGAACAGTGGCGAGATATATTATCAAGGTGTCATCTTAATCATACATGA	1499
OY	1621	CCCAAGGCGATGCCAGACATTAATCTCTGTGAAAGCTATGACATCAAGGTAATTACTG	1680

```

Db      1499 GTTCTG---TGCCGGAATTTGCGGCAGATCGCAGAAAGCTTATGTCATGTTGGATCCGTA 1555
QY      1681 TGCACAAAGCGGGAAGATTTGGCCCCCGGCGAGTCGCCGAATAGCTAGCCACAGTGGTCTG 1740
Db      1556 TTCTCTCACCCCTGAGTAGAGCTGCAAGCTGGGCTGAAAAAGCGTTAGCGCATGAAAAGATCGTC 1615
QY      1741 TGGTG---ATGAGATGTGTGTGTCAAAAAAGTGAATACTGTTACCTTATGATTGCCCCCG 1797
Db      1616 TGGTGTTTGTGCAATTAATGATGTCATGAAACTGAGCAGCTTACCAATGCAATCAAG 1675
QY      1798 GCATAGTAATGCCCAATG 1817
Db      1676 CGGAAGGAGATGACAAATG 1655

RESULT 8
US-10-282-122A-11812
; Sequence 11812, Application US/10282122A
; Publication No. US20040029123A1
GENERAL INFORMATION:
APPLICANT: Wang, Hiansu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trewick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/0282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11812
LENGTH: 1875
TYPE: DNA
ORGANISM: Burkholderia cepacia
US-10-282-122A-11812

Query Match      19.3%; Score 378.8; DB 13; Length 1875;
Best Local Similarity 54.7%; Prd. No. 7,7e-116;
Matches 938; Conservative 0; Mismatches 747; Indels 31; Gaps 8;

136 ACCGGGGCTTAATACCTGATGATAGCTGAAGAAAGCCCATGGGTCATAACACATTTTGGC 195
Db      178 ATCGGCGGCGACCGGTGCTCATGAAGAGCACTGGCCGACGAAACGTCGATTCATCTGGGGC 237

```

196 TATCCCGCGGCGCAATTTTGGCCCATCTATGATGAACCTGTACCGCTTTGAAAGCGGCGGG 255
238 TACCCCGCGGCGCTGGTTCTCTACATCTAGACGAGCTGTACA-----AGCAGGAC 288
256 GAAATTTAGCATATTTTGGTGGCCATGAAGAGAGCTTCCATGCGGGGATGGGTAT 315
289 AAGATTGACACGCTGGTGGCCATGAAGAGCGGCGCTGTACGAGCGAGTGGTAT 348
316 GCCAGAGCCAGAGTAAAGTGGAGTTTGTTCGCTACATCTGGAACGAGGGGCGACTAC 375
349 GCGGCTCCACCGGCAAGTGGCGTCTGCTCTGTGAGCTGGGGCCCCGGTGTACCAAT 408
376 TTGGTACCGGCAATTCGCAATGCGCATTTGGACTCGGTCGCCCATGTGTGATTACTGA 435
409 GCGGTGACCGGCAATGCGCAACGCGCTTACATGATTCGATCCGATGTGTATCAGCGGC 468
436 GAGGTGGCGCGTGCATGATGTGTAGAGATCTTTCCAGAGAAATGACATTTTGGCATC 495
469 CAGGTGCGACCGCGCGATCGGTCAAGATCGCTTCCAGAGATCGCACCGTGGCATC 528
496 ACCTTACCGATGTTAAGACACTCTATGTGTAGTGTAGTGGCGGATATGCTCGCAT 555
529 ACGGTCGCGTGTGAAGCAACCTCTCGTGAAGACGTCGCGCATCTCGCGGAAAC 588
556 GTTACTGAGGCTTCCATCTTGTAGACACCGGTGTCCCGGCGCGGTTTGTGATGATAT 615
589 GTCAAGAGCGCTTATATCGCCGACCGGCGGTGCGGCGGCTGTGTATGATC 648
616 CCCAAGATGTGGGCTTAAAGAAATGTGATGATATCCCTCGACCCCGGTGACCTTAT 675
649 CCGAAGGACATCTGGAAGACCGGTGCGAGTACGACGATCA-----GAGGTGTG 702
676 CTACCGGCTTATCGCCCGCGGTTAAAGTATCCCGGCAAAATTAATGCGGATGCA 735
703 CTGGTTGTGTAACACCCCGTACAGAAAGGCGCATTCGAGTCCGAAAGCGGTGTG 762
736 TTGTTGAGACGAGCGCAAAATCCCTGTCTACGTAGAGGAGGAGGATGCGCCCAT 795
763 CTGCTGTGACGAGGGAACCGTCCGTACATCAACGAGGCGCGCATCTCTGCGGAC 822
796 GCCCATGCCAGGTGCAAGAAATTTGCGAAAGGTTTCCAGTTGCGGTTAAACAACCTG 855
823 GCGTGCCTGTGAACGACGTTGCGGACCTGTCTGCTACACCGGTGCAAGACCTG 882
856 ATGGGAATTTGGGGCTTTGACGAAACCATCCCTTTCGAGGATGTTGGGATCAT 915
883 ATGGGCTTCGCGGCTATCGCGCTCGGACAAAGAAATTCCTCGGCACTGCTGCGATC 942
916 GCGCACCGCTATGCGCACTTTCGCGTCAAGCATGTGATTTGTTGATGAGTGGGCGC 975
943 GCGACCTTACGAAAGCAATGCGCATGCGACATGCGACGTCGTGATCGCGATCGGCGC 1002
976 GCTTTCGACGACCGGTTACTGCAACTAGCAATTTGCTACCGCGC---CAAGTA 1032
1003 GCTTCGACGACCGGTGATCGGCGACCGGCGCATTCCTGCTGCGCGCGCGCAATC 1062
1033 ATTCAATTGACATGACCCGCGCGAGGTGGGAAAAAACAAGGCTCCCGATGTGCCAT 1092
1063 ATTCAATGACATGACCCGCTGCTGATCTGAAAGCGGTGAAGTGTGACATCCGATC 1122
1093 GTGGGAGTGTACCGCATGTTTGAACAGCTTTTGAAGCGGCGCGCGGAATGATTA 1152
1123 GTCCGCGAGTGAAGAAAGTGTCTAAGAGCTGATGACGAGTTGCGACGCGCGACAT 1182
1153 CCCACCATTCCTTACCATCCAGGCAAGCATGTTAATGACATGATGATGCGCGACAT 1212
1183 GGGCCCGACACCGAGCGCTGCGCATGTGTGAAGACATCGAAGGCTGGCGCGAAG 1242
1213 TACCCCTTCAGGTGCGCCCATATGAGGATCTTTTGGCCCCCGAGAGTATACAGAA 1272
1243 GACTGCTTGAAGTACGACCGGAAAGCGAGATCATCAAGCGCGAGTGTGTGAGAG 1302
1273 ATTGGTGCAGGCGCGCG---ATGCTACTACACCAACGATGTGGAGCAACACCAATG 1329

1303 GCGTGGAGCTGACCGAGCGGCAATGCGTTCGTGTGCTGTGACGCTCGGCAACACAGATG 1362
1330 TGGGCGGCGCAGTTT---TGAACATGAGCCCGCGCATGATGATTTCCAGTGTGCTTG 1386
1363 TGGGCGGCGCAGTTTACCGTTTCAACAGCGCGGTGCTGTGATCACTCGGCGGCTC 1422
1387 GGTACGATGGGCTTGTGTTTACCTGCGCCATGAGAGCGCAAGTGGAGTGGAGCA 1445
1423 GGCACGATGGGCTTGGCTGCGCGCGCGATGAGCGTCAAGATGAGCCCATCCGACAC 1482
1446 GGGGTCAATTTGATGATGAGGAGATGCCAGCTTCCAAATGATCTTCAGAACTGGAAAC 1505
1483 GAGGTGTGTGATACGAGGAGAGGCTGATTCAGATGTGATCTCAGAACTGTGACAC 1542
1506 CTAGCCAGTACGACATCCAGGTTTAACTATTTCTCAATTAACGTTGGCAGGAGATG 1565
1543 TGCCTGAGTACGACACCGCGGTGAAGATCAATTCGTGAACACCGCTACCTCGCATG 1602
1566 GTGCTGTGTGCAACAACTTTTCAAGAAAGCTTATCTGCTTCTTAAGATGCTCCAG 1625
1603 GTCCCGCAGTGCACAGATCGAAATACGACGAGCGCTATTTGCAATTCGTACATG--GAT 1659
1626 GGCATGCCAGACATTAATCTCTCTGTGAGCTTATGACATCAAGGATTAATCTGTGCGC 1685
1660 GCGCTGCGGANTTTGTGAAGCTGCGGACGTAAGCGCATGTGCGCATCGGATCGAA 1719
1686 AAGCGGAAAGATTGGCCCGCGCATGCGCGAATGCTAGCCCAATG--GTCTGTG 1742
1720 AAGCTTGGATGTGAGCCCGCGCTGAGAGAGCGCTGCGCTGAAAGACCGACCTG 1779
1743 GTGATGATGTGTGTGCAAAAAGATGAAACTGTTCCTTATGATTTGCCCGCGCATG 1802
1780 TTTCGTGACTTCCAGACCGATCCGACCGAAACGTCGTGCGGATGTGACAGCGCGCAG 1839
1803 AGTAATGCCCAATCTAGGTTTACCGGAAGTGGC 1838
1840 GGCATCCAGAGATGCTGCTGATCGGAAGATCG 1875

RESULT 9
US-09-815-242-7947
Sequence 7947, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7947
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1725)
; US-09-815-242-7947

```

```

Query Match      19.1%; Score 373.8; DB 9; Length 1725;
Best Local Similarity 54.6%; Pred. No. 3.5e-114;
Matches 886; Conservative 0; Mismatches 712; Indels 25; Gaps 6;

```

```

QY 135 AACGGGCGCTTATCTGATGATGAGCTGAAAGCGCATGGGGTCMAACACATTTTGG 194
DB 12 ATCTGGCGCTGAATGCTGCTCCGCTCGTTGGCGACAGAGCGTTAAGTACATCTACGG 71
QY 195 CTATCCCGCGGGCGCAATTTTCCCATCTATGATGAACCTGTACCGCTTTGAAGCGCGGG 254
DB 72 GTACCCGGGTGGTGGCTCTCTCATATCTACGACCGCTCTTC-----AAGGAGCA 122
QY 255 GGAATTTGACATATTTTGTGGCGCATGAACAAGAGCTTCCCATGGGGCGATGGGTA 314
DB 123 CGACGTGACCCACATCTGCTGGCGCAAGACAGCGGCTACCCATATGCGCGAGCTA 182
QY 315 TGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGCTACATCTGAGCCAGGGCGACTAA 374
DB 183 CGCCCGCGCACCGGCAAGCCGGTGTGCTGTGTACTTCGCGCGGGGGGAGCA 242
QY 375 CTGTGTGACCGGCATTTGCCAATGCGCATTTTGGACTCGTGGCCATGTGTGATTTCTGG 434
DB 243 CGCCATCACCGGTATCGCCACCGCATGACACTCCATCCGATGTGATCTGTCCGG 302
QY 435 AGAGTGGGCGGTGCGCATGATGTGAGCGATCTTCCAGAGAAATTCATTTTGGCAT 494
DB 303 CGAGGTGCGGACCAATGCTGTGATCGATGCTTCCAGAGAAACCGCATGTGCGGAT 362
QY 495 CACCTTACCGATCTTAAAGCACTCTTATGTGTGATGATGAGTGGCGGATATGCTCGCAT 554
DB 363 CTCCCGGCGCATGTGAAGACAGCTTCATCATCAAGCATCCCTCGGAAATTCGCGAGT 422
QY 555 TGTATCTGAGGCTTCCATCTTGTAGACACCGGTGCTCCCGGCGCGTTTGTGATCTGAT 614
DB 423 GATCAGAGAGCGCTTCTTACTGCGCGCATGCGGTGCTCGGCGCGGCGGTGCTGTGAT 482
QY 615 TCCCAAGAGTGTGGCTTAAAGAAATGTGAGTACATTCCTCGACCCCGGTGACGTAA 674
DB 483 TCCGAGAGACATGGGCGACCGGACGCAAGAAATTCGATATTCCTTCGAGAGAGGTCA 542
QY 675 TCTACCGGGTATCGCCCCACGGTTAAAGTATCCCGACAAATTAATGGGCAATGCA 734
DB 543 GTTGGCTTGTACACCGCGCGCTTGTGCGGTCACTCGGAGACATCCGAAAGCGCGCGA 602
QY 735 ATTGTGAGAGCAGCAGAAATCCCTTCTACGTAGGAGGAGGCGCATGCGCGCAA 794
DB 603 GATGCTCTGCGCGCGCAAGCGCGCGGTGTGTATTCGCGCGCGCGGTGATCATGCGGAA 662
QY 795 TGCCATGCGCGCGGTGCGAGAAATTTGGCGAAAGTTTCCAGTTTCCCGGTAAACAACCT 854
DB 663 TGCCCGCGCGCGGTGCGAGCGGTGCGAGTGTCTCAACCTGCGCGGTGACCAACCTCT 722
QY 855 GATGGGAATTTGGGCTTTTGAAGAAACCATCCCTTCCGATGGGTATGTTGGGATGCA 914
DB 723 GATGGGCTCGCGCGGTATCCCGCGCGAGACCGCATGTTCTCGGATGCTCGGAGATCA 782
QY 915 TGCCACCGCTATGCGCAACTTGTCCGTCAAGCATGTGATTTGTGATGCAAGTGGGCGC 974
DB 783 TGCCAGCTTCAACCGCAACTGCGCATCATCAACGACGATGATCTCGCGGTGCGGCGC 842
QY 975 CGGTTTCAAGCAGCGGCTTAACTGCGCAACTGAGAGATTTTGTGACCGCGCGCAAGTAT 1034
DB 843 GCGTTTCAAGCAGCGGCTTCAACGATGCGGCGCAAGTTTGTGCGCGCAAGCGCAAGTAT 902

```

```

QY 1035 TCACATTCATGACATGACCCCGCGAGGTGGGAAAAAACAAGGCTCCCGATGTGCCATTTGT 1094
DB 903 CCACATGACATGATGATCCGCGGTGATTTCCAAAGACATCAAGGCGGACATCCCGATCGT 962
QY 1095 GGGGATGTAGCGCATGTTTAAAGACGCTTTTGCAGCGGGCGCGGAAATTTGATTAACC 1154
DB 963 CGGCGCGGTGAGAGGCTTCTCAACGAAATGTGTGATGTGAGGAAATTCGCGGAGAC 1022
QY 1155 CACCATTCCTCCATACCAACCGCATGTGTTAAATTCGATGATCATTTGCG--GACCGA 1211
DB 1023 CCGGAAACAGAGTCCCAAGGCTCTGTGTGAAGCAGATGACGATGTGGGTGTAAACCG 1082
QY 1212 TTACCCCTCCAGGTGCCCCACTATGAGTACTATTTGCCCGCCAGAGGTAGTACACGA 1271
DB 1083 TGGGTGTTCCGTAGACAAAGGCGACGCGACATCATCAAGCGGACGCTCATCGA 1142
QY 1272 AATTGTGCGCAGGCCCC-----CGATGCTTACTACACACCGATGTGGGACACACA 1325
DB 1143 GACGCTTACAGATTAACCATGCGATGCTTCTATCACTCCGATGTGCGGACAGACA 1202
QY 1326 AATGTGGGCGGCCAGTTTGA---CAATGCCCCCGCGCATGTGATTTCCAGTCTGG 1382
DB 1203 GATGTTGCTCTGCACTACTACAAGTTCAACAAAGCCCAATGCTGTGATCAATCCGCGG 1262
QY 1383 CTGGGTACGATGGGCTTTGTGTTTACCTGCGCCATGAGGAGCCAAAGTGGAGTGGGGA 1442
DB 1263 TCTGGACCATGAGGCTTGGCTTCCCGGCGCGGATGGGATCAAGCTCAACTCCCGGA 1322
QY 1443 CGA-GCGGTATTTGCATCAGTGAAGATGCGCACTTCCAAATGATCTTCAGGAATCGG 1501
DB 1323 CGAGATGTGCTGTGTGACCGCGGAGGCGAGTATCCAGTAAATCCAGAACTGTG 1382
QY 1502 AACCTTAGCCAGTACGACATCCAGGTTAAACTTATCTCAATAACGTTGGGAGGG 1561
DB 1383 TACCTGCTGCAATGACACTGCGGTGAAGATGCTCAACTGAACAATGTGCTCCTGGG 1442
QY 1562 GATGTGTGCTGATGCGCAAACTTTCTAGAGAAAGCTTATGCTTCTTCAATGTG 1621
DB 1443 CATGTGCGCGCAATGCGAGACATGCAATCAACACCGTTATTCGATTCCTCAATG-- 1500
QY 1622 CCAAGGATGCGCAATTAATCTCCTGTGAGAGCTATGAGATCAAGGTTATCTGT 1681
DB 1501 -GATGCTGCGGAACTTGTCAAGCTTGTGCGGAGCTTACGAGCATGTGCGGAT 1559
QY 1682 GCGCAAGCGGAGATTTGGCCCCCGCGATGCGCGAATGCTAGCCCAATGTGCTGT 1741
DB 1560 CACCGACCTGAAGATCTGAAGCGAAGATGAGGAGGCGTTGGCATGAAGATGCGCT 1619
QY 1742 GGT 1744
DB 1620 GGT 1622

```

```

RESULT 10
US-10-282-122A-30524
; Sequence 30524, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Malone, Cheryl
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Car, Grant
; APPLICANT: Trawick, John
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITPA.034A

```

```

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 30524
LENGTH: 1725
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-30524

Query Match      19.1%; Score 373.8; DB 13; Length 1725;
Best Local Similarity 54.6%; Pred. No. 3.5e-114;
Matches 886; Conservative 0; Mismatches 712; Indels 25; Gaps 6;

QY 135 AACGGGGCTATATCCGATGATAGCTGAAAGCGCATGGGGTCAACACATTTTGG 194
DB 12 ATCTGGCGCTAAATGTCGTCGCTCGTTCGGGAGAAAGCCTTAAATCACTACGG 71
QY 195 CTATCCGCGGGGCAATTTTGCCTATATGATGAACTGACCGCTTTGAAGCGCGG 254
DB 72 GTACCGGGGTGTGCGCTCTGCAATCTACGACGCGCTTC-----AAGAGCA 122
QY 255 GGAATTTAGATATTTTGTGGCCGCAAGAAAGAGGCTTCCGATCGCGGATGTGTA 314
DB 123 CGAGGTGACCAATCTCGGTGCGCCAGAAAGCGGCTAACATATGCGCGACGCTA 182
QY 315 TGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGATACATCTGACACAGGCGACTAA 374
DB 183 CGCCCGGCGCAACCGGCAAGCCCGGTGTGTGTGATCTTCGGCCCGGCGGACCA 242
QY 375 CTGTGTGACCGGCAATGCGCATTTGCACTCGGTGCCCATGTGTGTGTGATTTGG 434
DB 243 CGCATACCGGTATCGCCACCGCATATAGACTCCATCCGATGTGATCTGTGGG 302
QY 435 AGAGGTGGCGGTGCGCATGATGTAGCATGCTTTCCAGAAATTTTGTGAT 494
DB 303 CCAAGTGGCGGCAACATGTCGTGATCGGATCGGCTTCAGAAACCAATGTCGTGGAT 362
QY 495 CACTTACCGATCGTTAAGCACTCTATGTGTAGTATGTCGCGGATATGTGCTGCAT 554
DB 353 CTCCGCGCGCATCGTGAAGCAAGCTTCATCATCAAGCATCCCTCGGAAATTTCCGAGT 422
QY 555 TGTATCTGAGCTTTCATCTTGTGTAGACCGGTGTCCCGGCGGCTTTTGTATCAT 614
DB 423 GATCAAGAGGCTTTCTTCTGTGGGCACTCGGTGTCTCGGCGGCTGTGTGTAT 482
QY 615 TCCAGAGATGTGGCTTTAAGAAATGTAGTATTCCTCGACCCCGGTGACGTAA 674
DB 483 TCCGAAGACATGGCGGACCGGACGAGAAATTCGAATATCTCATCGAAGAGTCAA 542
QY 675 TCTACCGGTTATCGCCCAAGTTAAAGTAAATCCCGACAAATTAATGCGGATTCGA 734
DB 543 GTTGGCTGTGACAGCCCGGCGGTTCGCGGTACTCGGAGACAGATCCGCAAGCGCGCA 602
QY 735 ATTTGTGAGAGCGCCAGAAATCCCTTGTCTACGTAGGGGAGGGGATCGCCGCCA 794
DB 603 GATGCTCTTGCGCGCAAGCGCCGCTGTATTTCCGCGCGCGCGTATCATAGGGCA 662
QY 795 TCCCATGCCAGGTGCGAGATTTTGCAGAAAGTTTCCAGTTTGCCTGATACACACCT 854
DB 663 TCCCGCGGCGCGGTGACCGAGCTGCGCAGATGTGTCACTGCGCGGATCAACACCT 722
QY 855 GATGGAAATTTGGGGCTTTTGAAGAAACATTCCTTTCGATGGGTATGTTGGTATGCA 914
DB 722 GATGGGCTTGGCGGGTATCCCGGCGAGACCGCATTCCTCGCATGTCGCGGATGCA 782
QY 915 TGGCACCGCTATGCGCACTTGTGCGGTACGAAATGTATTTGTATGTCAGTGGGCGC 974
DB 783 TGGCAGCTTTCACCGCCAACTGCGCATCATCATCAAGCGACGATTCCTCGCGTGGCGC 842
QY 975 CCGTTTCAAGACCGGGTACTGCGAAATGACGAATTTCTCTACCGCGCAAGTAAAT 1034
DB 843 GCGTTTCAAGACCGGGTACTGCGAAATGACGAATTTCTCTACCGCGCAAGTAAAT 902
QY 1035 TCACATTCACATGACCGCGCGGAGTGGGAAAAACAGGGCTCCCGATGTGCCATGT 1094
DB 903 CCACATTCACATGACCGCGCGTTCGATTCGAAAGACATCAAGCGCCGACATCCGATGT 962
QY 1095 GGGGAGATGACCGCATGTTTGAACAGACTTTTGCAGCGGCGCGGAAATGATACCC 1154
DB 963 CGGCGCGGTGACAGCGTTCTTCAACGAATGATGCGATTCGTCAGGGAAATGCGCGAGC 1022
QY 1155 CACCCATCCCATACACCCGAGCATGTTTAAATGCGATGATATGCGC--GGAACGA 1211
DB 1023 CCGGAACAGAGTGGCCAGGCTGCTGTGGAAGATGACGAGTGTGCTGTATACG 1082
QY 1212 TTACCCCTTCAGGTGGCCCATATAGAGATATATGCCCCCGAGAGATGATACAGA 1271
DB 1083 TGGGTGTTCCCGTACGAAAGGCGAGCGAGATCATCAAGCGCGACCGTATCGA 1142
QY 1272 AATTTGGTCCAGGCCCC--GATGCTTACTACACACCGATGTGGACAAACACA 1325
DB 1143 GACGCTTACGAGGTATACCATGCGGATGCTTATCATCTCGATGTGCGCACACCA 1202
QY 1326 AATGTGGGCGGCGCAATTTTGA--CAATGCCCCCGCGAGATGATTTCCAGTGTG 1382
DB 1203 GATGTTCCGCTGCGATCTACATGATTCACAGGCCAATGCTGTGATCACTCGGCG 1262
QY 1383 CTGGGTACGATGGGCTTTGTTTACTGCGCCATGCGGACCAAGTGGAGTGGGGA 1442
DB 1263 TCTGCGCACATGGGCTTCGCTTCGCGCGCGATGCGATCAAGCTCAACTTCCGGA 1322
QY 1443 CGAGCATTTGATCATCAATGAGATGCGAGCTTCCAAATGAAATTTTCAGAACTGG 1501
DB 1323 CGACATGTCGCTCTGTCGACCGGCGAGGCGAGTATCCAGATGAATTCAGAACTGTC 1382
QY 1502 AACCTTACCGCATGACATGACATGAGTTAAATCTATTTCTCAATACGTTGGCAGG 1561
DB 1383 TACTGTGCTGAGTACGACTGCTCCGCTGAAGTCTCACTGAACAAATGTCGCTGG 1442
QY 1562 GATGTGCTGCTGAGGCAACAACTTTCTAGAAAGAGTATCTGCTTCAATGTC 1621
DB 1443 CATGTGCGCAGTGCAGACATGACAGTACAAAGCGCTTATTCGATTCCTACATG-- 1500
QY 1622 CAGGGCATGCGACATTTATCTCTGTGAAGCTTATGAGCATCAAGGATATTACTGT 1681
DB 1501 -GATTCCTGCGGACTTGTCAAGCTTGCAGAGCTTACGCGGACATGTGCGATGCGCAT 1559
QY 1682 GCGCAAGCGGAAATTTGGCCCGGCGCATGCGGAAATGCTAGCCCAATGTGCTGT 1741
DB 1560 CACGACCTGAAGAGATCTGAAGCCGAAATGAGAGGCGTTTGGCATGAAGAAATGCGCT 1619
QY 1742 GGT 1744
DB 1620 GGT 1622
```

RESULT 11
US-09-815-242-5925
Sequence 5925, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Cair, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5925
LENGTH: 1815
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1815)
US-09-815-242-5925

Query Match 18.8%; Score 368.6; DB 9; Length 1815;
Best Local Similarity 55.3%; Pred. No. 2,1e-112;
Matches 805; Conservative 0; Mismatches 634; Indels 16; Gaps 4;

QY 172 CATGGGGTCAAAACATATTTTGGCTATCCCGCGGGCAATTTGGCCATCTATGATGA 231
DB 139 CAGGGCGTTAAACAAGTATTCGGTTATCCCGAGCGCGATCTTGATATTTATGATGCA 198
QY 232 CTGTACCGCTTTGAAGCGGCGGGGAATGAGCATATTTTGGTGGCCATGAACAAGA 291
DB 199 -----TTGCATACCGTGGGTGTATGTATCATGTATATGTTTCCATGAGCAGCG 249
QY 292 GCTTCCCATGGCGCGGATGGGTATGCCAGAGCCACAGTAAAGTGGAGTTTGTTCGGT 351
DB 250 GCGGTGATATGGCGGATGGCTTGCGCGCGCGACCGGGGAAGTCGGGTGTCTGTGTA 309
QY 352 ACATCTGAGCCAGGGGCGAGCTAATTGTGACCGGCAATGGCCATTTGCCATTTGGACTCG 411
DB 310 ACGTGGGTCCAGGGGCCAAGCCATATCTGCAATCGCCACCGCTTATATGATATCC 369
QY 412 GTGCCCATGTGTGATTAATGAGAGGTGGGCGGTGCCATGATTTGATGCGATGCTTTC 471
DB 370 ATTCATTAGTTGTCTTTCCGGGCGAGGTAGCACTGTTGATATGTTAGTACGCTTT 429
QY 472 CAGGAATATGACATTTTGGCATACCTTACCGATCGTTAGCACTCTCTATGTGTACGT 531
DB 430 CAGGAGTCCGACATGTGTGGGATTTTCCGACCGGTGTAAACAACAGTTTCTGTGTAAG 489

QY 532 AGTCCGGCGATATGCTCGCATTTGTACTGAGGCTTTCCATTTGTGACACCGGTCGT 591
DB 490 CAACCGAACAATTCGCGAGGTGTGAAAAAGGCTTTCTGGCTGGCGCAATGTGTCCG 549
QY 592 CCCGGCGCGTTTGTATGATATTTCCAGAGATGTGGGCTTAGAAGATGTGATCAT 651
DB 550 CCAGAACCAATGATCTGTGATTTTACCGAAGATATTTCTTATATCCGGGAACAAATGCC 609
QY 652 CCCCTGACCCCGGTGACGTTAATCTACCGGTTATCGCCACGGTTAAAGTTATCCC 711
DB 610 TATGTCGGCGGAGTCGATCAGTATGCTTTTAAATCCCACTACTACCGGACATTA 669
QY 712 CGACAAATTAATCGGATTTGCAATTTGTGACAGCCAGAAATCCCTTGCTCTACGTA 771
DB 670 GGGCAATTAAGGTGTCTGTGCAACGCTGTAGCGGCAAAAAACCGGTTGTCTACGTA 729
QY 772 GGGGAGGGGCGATCGCCCGCAATGCCATGCCCCAGGTGACAGAAATTTGCGAAGGTTCC 831
DB 730 GCGGTGGGCAATCAAGCGGGCTGCGATCAGCAATGTAAGAAACGGTGAAGGGGTTG 789
QY 832 CAGTTGCGGTATCAACCACTGATGGGAATTTGGGCTTTTGAAGAAACATTCCTT 891
DB 790 AATCTGCCGTGTGTCTCATTTGATGGGCTGGGGCGTTCCGGCAACGATGCTCAG 849
QY 892 TCGGTGGATATGTTGGGTATGATGACCAACCGCTATGCCAATTTGCCGTCAAGCATG 951
DB 850 GCACTGGGCAATGCTGGGAATGACAGTACTTACGAAGCCAAATGACATGATCAACGG 909
QY 952 GATTTGTTATGAGTGGGGGCGCGTTTGCAGCAGCCGGTAACTGGCAAACTTAGCA 1011
DB 910 GATGTGATTTCCCGGTGGGATACGATTTGATGACCGAAGCAACAACTGTGCAAG 969
QY 1012 TTGCTAGACCGGCCAAAGTAATTCATTTGATGATGACCGCGGAGGTTGGGAAAAAC 1071
DB 970 TACTGCCCAATATCCACTGTCTTGTCAATGATGATATGATCTTCAATTTCTTAAC 1029
QY 1072 AGGCTTCGCGATGCGCATTTGTGGGGATGTAAGCCATGTTTGAAGACGCTTTTGCAG 1131
DB 1030 GTACTGCGGATATCCGATGTTGGGGATGCTGCCAGTCTCTCAACAAATGCTTGA 1089
QY 1132 CCGGCGCGGATATGATTAACCCCACTCCCATACCAACCCAGGCAATGTTAAATCCG 1191
DB 1090 CTCTTGTGCAAAATCCGCGCATCAACCACTGATGAGATCCGCACTGTGTGCGACA 1149
QY 1192 ATTGATCATTTGCGGACCAATTAACCCCTCCAGGTGCCCACTATGAGATATATTGCG 1251
DB 1150 ATTGAACAGTGGGCGCTGTGCAGTGCTGAAATATGACATCAAGTAAAAAGTTAA 1209
QY 1252 CCCGAGAGGTAGTACACGAATTTGTCGCAAGGCC---CGATGCTTACTACACACC 1308
DB 1210 CCGCAGCGGTGATTCAGACTCTTTGCGGATTTGACCAAGGAGACCTTACGATGAGTC 1269
QY 1309 GATGTGGGCAACACCAATATGTCGGCGCC---AGTTTGAACAATGAGCCCGCCGCA 1365
DB 1270 GATGTGGGCAACCAAGATGTTGTCGACTTTATATTCATTCGACAAACCGGCTCCG 1329
QY 1366 TGGATTTCAATGCTGCTTGGGTAGATGAGGCTTTTGAACCTGCGCCATGAGAGCC 1425
DB 1330 TGGATCAATTCGAGTGGCTCGGCAAGATGAGTTGTTTACCTGCGGCACTGGGCGTC 1389
QY 1426 AAGTGGAGTGGGGGAGCAG--CGGCAATTTGATAGTGGAGATGCCAGCTTCCAAAG 1484
DB 1390 AATATGCGTTGCCAAGAAACCGTGTGTGATCACTGGGCAAGATTCAGATG 1449
QY 1485 AATCTTCAGAACTGGAAACCTTAGCCAGTACGACATCCAGATTAAACTTATTTCTC 1544
DB 1450 AATCTTCAGAACTGCTACCGGTTGCAATACGATTCGCCGTATGATGATATC 1509
QY 1545 AATTAAGTGTGGAGGAGTGTGTGCTCAGTGGCAACAACTTTCTAGCAAGAGTTAT 1604
DB 1510 AATTAAGTGTGTGGAGTGTGAGCAGTGGCAGACATATTTCCGGCGTCA 1569
QY 1605 TCTGCTTCAAGATG 1619

DB 1570 TCAGATCTTATATG 1584

RESULT 12

US-10-282-122A-20222
Sequence 20222, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20222
LENGTH: 1815
TYPE: DNA
ORGANISM: Escherichia coli
US-10-282-122A-20222

Query Match 18.8%; Score 368.6; DB 13; Length 1815;
Best Local Similarity 55.3%; Pred. No. 2,1e-112;
Matches 805; Conservative 0; Mismatches 634; Indels 16; Gaps 4;

QY 172 CATGGGCTCAACCATTTTGGCATCCCGCGGGCAATTTGGCCATCTATGAGAA 231
DB 139 CAGGCGTTAAACAATTCGATTATCCGAGCGCAGTCTTGATATTATATGCA 198
QY 232 CTGTACCGCTTGAAGCGCGGGGAAATTGAGCAATTTGTGCGCCATGACAGAA 291
DB 199 -----TTGCAATACGTGGGTGATGATCATGATATGATTCGTCATGAGCAAGCG 249
QY 292 GCTTCCCATGGCGGATGGGTATCCAGACCCAGGTAAAGTGGAGTTTGTTCGT 351
DB 250 GCGGTGATATGGCGGATGGCTGGCGCGCGAGACCGGGAAGTGGCGTGTCTGTA 309
QY 352 ACATCTGAGACGAGGCGCACTAATTGTGTGACCGGCAATGGCCATTTGGAATCG 411
DB 310 AGGTGGGTCCAGGGGCGACCAATGCGATTACTGCGATCGCACCGCTTATATGATTC 369

QY 412 GTGCCATGATGATTAATCTAGAGAGGTGGGCGGTGCATGATTTGTAGCATCTTTC 471
DB 370 ATTCATTAATGATGCTCTTCCGGGAGGTAGCACTGTTGATATGTTAGATGCTTT 429
QY 472 CAGAAATTAATGATTTTGGCATCACTTACCATGTTTAAAGCACTCTTATGTGTACGT 531
DB 430 CAGAGTGGCATATGATGGGATTTTCGGACCGGTGTTAAACAGATTTCTGTGTAAG 489
QY 532 AGTGGCGGATATGAGCTCGCATGTTGTAAGGCTTTCATCTTGTAGCACCGGTGCT 591
DB 490 CAACGGAAGATATTCGCAAGTGTGAAAAAGGCTTCTGTGGCGGACAGTGTGTC 549
QY 550 CCAGACCAAGTATGATTTACGAAAGATATCTTAATCCGCGCAAAATTAACC 609
DB 652 CCCCTGACCCCGGAGATTAATCTACCGGTTATGCCCCAGGTTAAAGTAATCC 711
QY 610 TATGCTGGCGGATGCTGATGCTGATTAATCCCACTACTACCGGACATATA 669
DB 712 CGACAAATTAATGCGCATTTGCAATTTTGAAGCAGGCGCAAAATCCCTTCTTACGTA 771
QY 670 GGGCAATTAATGCGCTCTGCAACGCTGGTAGCGGCAAAAAACCGTTGTCTACGTA 729
DB 772 GGGGAGGCGCATGCGCGCCATGCTCCATGCGGAGGAGGATTTGGGAAAGCTTC 831
QY 730 GCGGTGGGCAATACGCGCGGCTGCATAGCATTTGAAGAAACGCTGAGGCGTTG 789
DB 832 CAGTTCGCGTAAACACCACTGATGGGAAATTTGGGCTTTTGAAGAAACATCCCTT 891
QY 790 AATTCGCGCTGTTTGTCTCATATGAGGCTGGGCGGCTTCCGGACACGATGTAG 849
DB 892 TCGGTGGATATGTTGGATATGATGCGCAACGCTATGCGCACTTTCCTGACGGAATGT 951
QY 850 GCACGTGGGATGCTGGGATGACCGGTACTACGAAGCAATATGACATCATTAACCG 909
DB 952 GATTTGTTGATGATGCGTGGGCGCGCTTTCGACGACCGGATTAATGCGAAACTAGACGA 1011
QY 910 GATGATATTTTGGCTGCGGCTGATGATTTGATGACCAAGCAAGATGATGCAAG 969
DB 1012 TTTGTAGCGCGCCAAAGTAATTCATTTGATGATGACCGCGGCGAGGTGGAAAC 1071
QY 970 TACTGCCCAATGCACTGTTCTGATTCATATGATTCATCTTCAATTTCTAAAC 1029
DB 1072 AGGCTCCGATGCGCATTTGGGGATGATGAGGCGCATGTTTGAACAGCTTTGCGAG 1131
QY 1030 GTGATCGGATATCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
DB 1132 CGGCGCGGGAATGATTTACCCCAACCATCCCATACACACGAGATGTTAATGCG 1191
QY 1090 CTCTGTGCAAGATTCGCGCATCAACCATGATGATGATGATGATGATGATGATG 1149
DB 1192 ATTTGATCATTTGGCGGATTAACCTCTCCAGGTGCCCATATGAGATATATTC 1251
QY 1150 ATTGAACAGTGGCGGCTGCTGATGCTGAAATATGACATCAAGTAAAGATTAA 1209
DB 1252 CCCAGAGTATGATCAAGAAATTTGTGCGCAGGCGC---CGATGCTTACTATCACAC 1308
QY 1210 CCGGAGCGGTATGAGACTCTTTGGCGGTGAGAGAGGAGACGCTTATGATGATG 1269
DB 1309 GATGTGGCAACCAACCAATTTGGCGGCGC---AGTTTGAACAAATGCGCGCGCA 1365
QY 1270 GATGTGGGACACACAGATTTGCTGCACTTATTTCAATTTGCAAAACCGGTGCG 1329
DB 1366 TGGATTTCAATGCTGCTGGGTGAGAGGAGCTTGTGTTTACCTGCGCCATGGAGCC 1425
QY 1330 TGGATCAATTTCCGATGCTGCGGACGATGAGTTTGTGTTTACCTGCGGCACTGGG 1389
DB 1426 AAAGTGGAGTGGGAGACGAG---CGGTCAATTTGATCAATGATGATGATGATGAT 1484
QY 1390 AAATATGCGTTCCAGAGAAACCGTGTGTTGCGTCACTGCGAGCGGAGATTTCAAGAT 1449

QY 1485 AATCTTCAGGAGTGGGAACCCCTAGCCAGTACGATCCAGCTTAAACTATTTATCTC 1544
 DB 1480 AACATTCAGGAGTGGTCTACCGCTTGCATACGAGTTGCCCTGATCTGATGATCTC 1509
 QY 1545 AATAACGCTTGGCAGGAGTGGTGGCTGAGTGGCAACAACCTTTCTACGAAGAAGCTTAT 1604
 DB 1510 AATAACGCTTATCTGGGAGTGGTGAAGAGTGGCAGGACATGATCTATTCGGCCGCTGAT 1569
 QY 1605 TCTGCTTCTTAACATG 1619
 DB 1570 TCACAATCTTATATG 1584

RESULT 13

US-10-282-122A-26894
 ; Sequence 26894, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; PRIOR FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-03-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See file Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 26894
 ; LENGTH: 1833
 ; TYPE: DNA
 ; ORGANISM: Moraxella catarrhalis
 US-10-282-122A-26894

Query Match 18.8%; Score 368.2; DB 13; Length 1833;
 Best Local Similarity 53.9%; Freq. No. 2.8e-112;
 Matches 908; Conservative 0; Mismatches 743; Indels 34; Gaps 6;

QY 137 CCGGGGCTTATATCTGTAGTAGTACCTGAAGCGCATGGGCTCAACACATTTTGGCT 196
 DB 119 CAGGGCTGAGAGTGGTCTTCAAGCATGTGATGAGAGTGTGAGATATTTTGGCT 178
 QY 197 ATCCGGGGGGAATTTTGGCATCTATGATGAATCTGTACCGCTTGAAGCGGGGG 256

DB 179 ATCCAGCGGTGGGATTTACATATTTATGATCTTTATTTAAGCAAGATA----- 229
 QY 257 AAATTGAGCATATTTGGTGGCCATGAAACAGAGCTTCCATGGCGGATGGGTATG 316
 DB 230 AAATTGAGCATATTTGGTGGCCATGAAACAGAGCTTCCATGGCGGATGGGTATG 289
 QY 317 CCAGGCAAGATTAAGTGGAGTTTGTTCGCTACATTTGATGACCAAGGGCCACTAAT 376
 DB 290 GTGCTACACAGCAGAGCGGGGTTTGGGCTTCCAGCTTCCAGGTCAGCAACCA 349
 QY 377 TGGTGAACCGCATTTGCAATGCCCATTTGGATCGGTGCGCATGTGTATCTAGAG 436
 DB 350 CTGTAAACAGCATTTGCAACTGCTTATATGATTCATCCCAATGGTTGTGTGGCAGT 409
 QY 437 AGGTGGGCGGTGCAATGATTTGGTATGATGATGATGATGATGATGATGATGATG 496
 DB 410 AGGTACCATCTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 469
 QY 497 CTTACCGATCGTTAAGCACTCCTATGTTAGTGTAGTGGGCGGATGGCTCGATG 556
 DB 470 CAGGCCCATCGTTAAGCACTCCTATGTTAGTGTAGTGGGCGGATGGCTCGATG 529
 QY 557 TTAAGAGCTTTCATCTTTGCTAGACACCGGTGTCGGGCGGCTTTGATGATGAT 616
 DB 530 TCAAAAAGCATTTTATATGTCAGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTG 589
 QY 617 CCAAGATGTGGGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 676
 DB 590 CAAAGATGATGACCAACCAAGTATGATGATGATGATGATGATGATGATGATGAT 649
 QY 677 TACCGGATATCGCCCGCAGGTTAAGTATCCCGACAAATTAAGCGCATTCAT 736
 DB 650 TCAGATCATATCAGCTTACCAATTAAGTATCCCGACAAATTAAGCGCATTCAT 709
 QY 737 TGTGAGAGCGGCAAGAAATCCCTTCTGATGATGATGATGATGATGATGATGAT 796
 DB 710 CGCTACTATCAGCAACCGCTGATGATGATGATGATGATGATGATGATGATGAT 769
 QY 797 CCATGCCAGGTGCAAGATTTGCGGAAGGTTGAGTGGCGGTAACCAACCTGTA 856
 DB 770 CCAATTAAGGCTGCTGATTTAGCGGATGATGATGATGATGATGATGATGATGAT 829
 QY 857 TGGGAATTGGGCTTTTGAAGAAACCATCCCTTTCGTTGGTATGTTGGTATGAT 916
 DB 830 TGGGACTGGGTGCTTATCCAGGCTCAGGCAACAATTTGGAAATGCTGATATGAT 889
 QY 917 GCCACGGCTATGCACTTTTGGCGTGGGAGGATGATGATGATGATGATGATGATGAT 976
 DB 890 GCAATATGAGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 949
 QY 977 GTTTCAGCAACCGGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1036
 DB 950 GTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1009
 QY 1037 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
 DB 1010 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
 QY 1097 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1156
 DB 1070 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
 QY 1157 CCGATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
 DB 1130 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
 QY 1209 -----CGATTCACCCCTCCAGGTGCCCATGATGATGATGATGATGATGATGAT 1261
 DB 1190 GTCTTGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
 QY 1262 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
 DB 1250 TGGTGAAGTGTCTTTTAAAGCTGACGATGATGATGATGATGATGATGATGATGAT 1309

QY 1319 AACACCAATGTGGGCGGCGCCAGTTT---TGAAACAATGGCCCCCGGCGATGATTTCCA 1375
DB 1310 AGCAATCAGATATTTTGTGTCATATTATTAAGATGAGCCACGCGCAATGGTTAACT 1369
QY 1376 GTGCTGCTGGGTACGATGGGCTTTGGTTTACCTGGCCGATGGAGCCAAAGTGGAG 1435
DB 1370 CAGGTGGCTTAGTACATAGGCGGTGGTTTACTTATGCTATGGAGCAAAATTTAGCTC 1429
QY 1436 TGGGGAGCAG- CGGTCAATTTGCATCAGTGGAGATGCCAGCTTCCAAATGAATCTTAG 1494
DB 1430 ACCCAAAAAGACAGTTGGTTGTATCACTGGCGAAGGCTCAATCAATTAACATCCAG 1489
QY 1495 AACTGGGAACCTTAGCCAGTACAGACATCCAGTTAAACTATATTCTCAATTAACGTT 1554
DB 1490 AGCTGTCACTTGTCTACAGTATATCTACAGATTAATTTAAATTTAAATTAATGCTC 1549
QY 1555 GCGCAGGGATGGTGGGTAGTGGCAACAATTCTTCTACGAAGACGTTATCTGCTTCTA 1614
DB 1550 AGCTTGGCATGTTAAGAGTGGCAGATGATGCTTTATGAAAGTGGCCATGCTCAACTT 1609
QY 1615 ACATGTCCAGGAGTGGCAGACATTAATCTCTGTGGAAGCTATGATGATCAAGGTA 1674
DB 1610 ATATG---CAATCTTTACTATTTTGTAAATTTGGAGAAAGTTATGGCATTAAGTG 1666
QY 1675 TTACTGTGCGCAGCGGAGATTTGGCCCGGCGATGCGCGAAATGCTAGCCCAATG 1734
DB 1667 TAAAAATTAACCAATCTGCTACATGCAAGAGAGCTAAACAGCCCTAGAAATGGATG 1726
QY 1735 GTCCGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1794
DB 1727 GTTGTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1786
QY 1795 CCGGC 1799
DB 1787 CTGGC 1791

RESULT 14
US-10-672-787-35
; Sequence 35, Application US/10672787
; Publication No. US20040067554A1
; GENERAL INFORMATION:
; APPLICANT: LAGACE, Robert, E.
; APPLICANT: PATTERSON, Robert, E.
; APPLICANT: BERG, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: ELITRA.025C1
; CURRENT APPLICATION NUMBER: US/10/672,787
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 09/596,002
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 96109
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-10-672-787-35

Query Match 18.8%; Score 368.2; DB 13; Length 96109;
Best Local Similarity 53.9%; Pred. No. 3.1e-111;
Matches 908; Conservative 0; Mismatches 743; Indels 34; Gaps 6;

QY 137 CCGGGCTTATATCTGATGATAGCCGTAACCCGATGGGTCAACACATTTTGGCT 196
DB 2433 CAGGGCTGAGATGCTTTGTAAGCATGATGATGATGATGATGATGATGATGATGATG 2492
QY 197 ATCCGCGCGGCAATTTTGGCCATCTATGATGATGATGATGATGATGATGATGATGATG 256
DB 2493 ATCCGCGCGGCGGATTTATCATATTTATGATGCTTTATTAAGCAAGATA----- 2543
QY 257 AATTTGACATATTTTGTGCGCCATGACAGAGAGCTTCCATGCGCGGATGGGTATG 316

DB 2544 AATTTGACATATTTTGTGCGCCATGACAGAGCGGCGCATATGCGCATGATGATGATG 2603
QY 317 CCAGAGCCACAGGTAAAGTGGAGTTTGTTCGGTACATCTGACACAGGGCGCATACT 376
DB 2604 GTGTACAAACAGCGAGAGCGGCGGTTTGTGGGCACTTCAAGTCCAGGTCAACCA 2663
QY 377 TGGTACCGGCAATTCGCAATGCCATTTGGAATCGGTCCCATGATGGTGAATTAAGAG 436
DB 2664 CTGTAAACGACATTTCACTGCTTATGATGATGATGATGATGATGATGATGATGATG 2723
QY 437 AGTGGCGCGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 496
DB 2724 AGTGTACATCTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2783
QY 497 CCTTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 556
DB 2784 CAGCGCCATCTTATGACAGATTTTCAAGTTGCGCATGCGGAGACATTTCCAAACAT 2843
QY 557 TTACTAGGCTTTTCAATCTTGTACACAGCGGTGTCGCGGCGGTTTGAATGATGATG 616
DB 2844 TCAAAAAGCATTTTATTTATGAGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCT 2903
QY 617 CCAAGATGTGGGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 676
DB 2904 CAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2963
QY 677 TACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 736
DB 2964 TCAGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3023
QY 737 TGTGAGAGCGGCGCAATTCCTTGTCTTACGTATGATGATGATGATGATGATGATGATG 796
DB 3024 CGCTACTATGACGCAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3083
QY 797 CCATGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 856
DB 3084 CCAAGTAAAGAGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3143
QY 857 TGGGAATTTGGGCTTTTGAACGAAACATCCCTTTCCGTGGGTATGATGATGATGATG 916
DB 3144 TGGGACTGGGTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3203
QY 917 GCCACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 976
DB 3204 GCACATATGACGAAACATGACATGATGATGATGATGATGATGATGATGATGATGATG 3263
QY 977 GTTTCAGCAGCGGCTTAACTGCAACATGATGATGATGATGATGATGATGATGATGATG 1036
DB 3264 GTTTGATGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3323
QY 1037 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096
DB 3324 ATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1156
QY 1097 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1166
DB 3384 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3443
QY 1157 CCCATGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
DB 3444 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3503
QY 1209 -----CGATTAACCCCTCCAGTGGCCCATGATGATGATGATGATGATGATGATG 1261
DB 3504 GTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3563
QY 1262 TAGTACAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1318
DB 3564 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3623
QY 1319 AACACCAATGTGGGCGGCGCCAGTTT---TGAAACAATGGCCCCCGGCGATGATTTCCA 1375

Db	3624	AGCATACAGATGTTTCCTGCATTTGTATTTAAAGTATGATGAGCAGCAATGGTTAACT	3683
Qy	1376	GTGCTGGCTTGGGTACGATGGGCTTTGGTTTACCTGGCCCATGGAGCCAAAGTGGAG	1435
Db	3664	CAGGTGGCTTAGGTACCAATGGGCGTTGGTTTACCTTATGCTATGGCAACAAATTAGCTC	3743
Qy	1436	TGGGGAGACGAG-CGGTCAATTTGCATCAGTGGAGATGCGAGCTTCCAAATGATCTTCAGG	1494
Db	3744	ACCCAAAAGACAGTTGTTTGTATCATCTGGGAGAGGCTCATCAATGAATCCACAG	3803
Qy	1435	AACTGGGAACCTTAGCCCAATGACATCCAGGTTAAACTATATTTCTCAATAACGTT	1554
Db	3804	AGCTGTCAACTTGTCTACAGTATTAATCAACAGTTAAATTTAAATTAATAATGCTC	3863
Qy	1555	GCGACGGGATGGTGGCGTAGTGGGCAACAACTTTAGAGAGAACGTTATCTGCTTCTA	1614
Db	3864	AGCTTGGCATGGTTAAGCATGGGAGATAGCTTTATGAAAGTGCGCATGCTCAATCTT	3923
Qy	1615	ACATGTCCAGGAGCATGCGACATTAATCTCTGTGAAGCCATGCGATCAAGGTA	1674
Db	3924	ATATG---CAATCTTTACTGTATTTGTTAAATTTGGCAGAAAGTTATGGCATTAAGGT	3980
Qy	1675	TTACTGTCCGAAGCGGGAAGATTTTGGCCCCGGCGATGCGCGAAATGCTAGCCACAATG	1734
Db	3961	TAAATATTAACCAATCTCTCTACCATGCAAGAAAGACTTAAACAGCCCTTAATAATGATG	4040
Qy	1735	GTCTGTGGTATGATGATGTGGTGGTCAAAAAAGATGAACCTGTTACCTTATGATTGGCC	1794
Db	4041	GTTTGGTATTTATTAATGATGTCTATGATGATTAAGGTAACATGTTTATCCGATGCAAGTGC	4100
Qy	1795	CCGCG 1799	
Db	4101	CTGG 4105	

```

/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 30934
/ LENGTH: 1725
/ TYPE: DNA
/ ORGANISM: Pasteurella multocida
US-10-282-122A-30934

Query Match      18.8%; Score 367.8; DB 13; Length 1725;
Best Local Similarity 54.8%; Pred. No. 3,7e-112;
Matches 821; Conservative 0; Mismatches 662; Indels 16; Gaps 4;

QY 128 AACGCCAACCCGGGGCTTATATCCTGATGATAGCTCGAAACGCCATGCGGCTCAACACA 187
DB 5 AAAAGTTATCTGGCGCGAGAAATGCTGTGCGAGCTCTTTCGCGATGAAAGGCGCTGAGTTT 64
QY 188 TTTTGGCTATCCCGGGCGGGGCAATTTTGGCCATCTATGATGAACTGTACCGCTTGAAG 247
DB 65 TATTTGGTATCTCGGCGGATCGGTTTGAATATTTATGATGCGATTCATACCTT----- 119
QY 248 CGGGCGGGGAATTTAGCATATTTTGGTGTGCGCCATGAAACAAGAGACTTCCCATCGCGCG 307
DB 120 ----GGGGGGGATTGAACATATCTTGTTCGACATGAAACAAGCGCGTGTGACATGCGAG 175
QY 308 ATGGGTATGCCAGAGCCACAGGTAAAGTGGAGTTTGTTCGGTACATCTTGACACAGGG 367
DB 176 ATGGTATATGCCAGAGCCACCGTAAAGGGGTGTGTCTTTGTGACATCCGACCCCGGTG 235
QY 368 CGACTAATCTGATACCGGCACTTCCCAATGCGCAATTTGACCTCGATGCGCATGAGTGTA 427
DB 236 CAACTAATACCATCACTGTGATTTGCAACAGCGTATATGATTCGCGCTTGTATATT 295
QY 428 TTACTGAGAGGTGGGCGGTGCCATGATGTGTAGCGATGCTTTCAGAAATTAACATTT 487
DB 296 TATCTGCAAGGTATGTACAGATTAATTTGCGATGATGCTTTCAAGAAATGCATATGG 355
QY 488 TTGGCATACCTTACCCGATGCTGTAGAGACTCTATGTGATGATAGTGGCGGAGATATNG 547
DB 356 TGGGTATCTCTCGTCCAGTCTCTAAACACAGTTTATATGTAATCTGCGGCTGAATTC 415
QY 548 CTCGATTTGTACTAGAGCTTTCATGCTGTGTACACCGTCTGCCGGGCGGTTTTGA 607
DB 416 CTGAACAATTAATAAAGCTTTTATATTTGGGTCAACAGGACACCGGACCTGTGGTGA 475
QY 608 TCGATATTTCCAAAGATGTGGGCTTAAGAATGTGATATTCCTCTGCACCCCGGTG 667
DB 476 TCGATTTGCCAAAGATGTGTATCTGTGACATAAATTAATTAATTAATCAAGAA 535
QY 668 ACGTTAATCTACCCGGTATATGCCGCCACGGTTAAGGTAAATCCCGCAAAATTAATGCG 727
DB 536 GAGTTGAATTAAGCTTTTATATATCAACAAATACAAAGGCAATAAGGCGAATTAATAAAG 595
QY 728 CATTCGATTTGTGAGACAGGCCAGAAATCCCTTCTCTACGTAGTGGGAGGGCGCATG 787
DB 596 CATTAATAAGCCTTATTAATGTGCGCAAAAAGCCGGTGTATTTGTGTGTGTGTGTATCA 655
QY 788 CCGCAATATGCCATAGCCACAGGTGACAGAAATTTGCGGAAAGGTTCCAGTGTGCGGTAAACA 847
DB 656 CTGGGATTTGAGGAGAAATTAACGCAATTTTGACAAACTTAATTTTGGCGGTCACTT 715
QY 848 CCACCTTGATGGGAATTTGGGGCTTTTGACGAAACCATCCCTTTTGGGTGATGTTGG 907
DB 716 CCTCTTAATATGGGATTTAGTGTGCTATCCAAAGCTCAGATTAAGCAATTTTAGGGGATGTTAG 775
QY 908 GTATGCAATGGCACCGGTATAGCAACTTTGCGCAAGCAATGTGATTTGTGATTTGCGAG 967
DB 776 GCAATGCAATGATCTTAACGAAGCAATATATGAAATGACATCATAGTATCTCATTTTGTGTA 835
QY 968 TGGGGGGCCGTTTTCAGACGCGGTTAATGCGCAAACTTAACGAATTTGTCTAGCCGCGCA 1027

```

Db 836 TTGGTGTGCGTTTGTGATGATGAAGACGACTAATTAAGTGGCAAAATATTTGTCCACATGCA 895
QY 1028 AAGTAATTACATTGACATGACCGCGGAGGTGGGAAAAACAGGGCTCCGATGTGC 1087
Db 896 AAGTGAATTCATATTGATGATTTGATCCAGCTCTATTCTAAAAATGTATTGGCATCTATTG 955
QY 1088 CCATTGTGGGGATGTACGCCATGTTTGAACAGCTTTTGCAGCGGGCCCGGGAATTGG 1147
Db 956 CAATTGTGGGAGTGCAAAAAAGTCAATTAATGAATTTTGAAGCTTATTAGAGATGACA 1015
QY 1148 ATTACCCCAACCATCCCATACCAACGAGCATGTTAAATCCGATTGATCATTTGGCGGA 1207
Db 1016 ATTATTTAAATTCGAAAATGCTTGGATGTGATGTGGACCAATGATGAATGGAAG 1075
QY 1208 CCGATTACCCCTCCAGGTGCCCACTATGAGATTAATTTGCCCCCAAGAGTATAC 1267
Db 1076 CAGTAAATGTCTCTCTTTTACGAGATGATTCGTCATTTAAACACAGCAAGTATGG 1135
QY 1268 ACGAAATTGTGCGCAGGCCCCCGAT--GCTTACTACACCAACGATGTGGACAAACCC 1324
Db 1136 AAGTGTATTATCGATTAAACACAGGTAGGCTTATGTGCTCAGATGTTGGACAAATC 1195
QY 1325 AAATGTGGGGCGGC--CCAGTTTGTAAACAATGACCCCGCGATGATTTCCAGTGTG 1381
Db 1196 AAATGTTGGGGCATTTACATTAATCCGTTTAAACAACCTCGGCGTTGATTAATCTGTG 1255
QY 1382 GCTTGGGTACGATGGGCTTTGTTTACCTGCGCCATGGAGCCAAAGTGGAGTGGGG 1441
Db 1256 GCTTAGGGACCATGGGCTTTGACTACAGCCGCACTTGGGTAAAGCTGCTTATCTG 1315
QY 1442 ACG-AGCGGTCAATTTGATCAGTGAATGCCAGCTTCAATGAATCTTCAGGAACCTG 1500
Db 1316 ACTCAACAGTGTGTGTGTGACGGGGATGGCAGTATTCANATGAATATTCAGAGCTTT 1375
QY 1501 GAACCTTAGCCCACTAGACATCCAGGTTAAACTATTATTCTCAATTAACGTTGGCAGG 1560
Db 1376 CTACTGCACAGCAATATGATTAATCCAAATGTGATTATCAGTTGAATTAACGCTTTAG 1435
QY 1561 GGATGTGCGTCAAGTGGCAACAACCTTCTACGAAGAAGCTTATCTGCTTAACATG 1619
Db 1436 GCATGTGAAAACAGTGGCAAGATCTGATTTATCTGTGAGCGCATTCACAGTCTTATATG 1494

Search completed: July 25, 2004, 13:28:24
Job time : 952.797 secs

maize cDNA sequences is either Virginia Malbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Source

Location/Qualifiers

1..2673

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:630902"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/Dupont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 13.5%; Score 265.4; DB 11; Length 2673;
 Best Local Similarity 50.9%; Pred. No. 1,26-65;
 Matches 757; Conservative 0; Mismatches 697; Indels 33; Gaps 8;

QY CCTAGCCACGCGAAACCGGGCTTATATCTGATGATAGCTGAAGCCCATGGGGTC 180
 DB CCCACGAGCCCCCGAGAGGGCGCGACATCTCTGTCAGTCCGACGCTCGGGGCTC 638
 QY AAAACATTTTGGCTATCCCGCGGGGCAATTTGGCCATCTATGATGACTGTACCG 240
 DB CGCGACGCTTCGCTACCCCGCGCGCGCTGCATGAGATCCACCGGACATCACCCGC 758
 QY TTTGAGCGCGGGGGAATTGAGCATATTTGGTGGCCATGAGCAAGAGCTTCCCAT 300
 DB TNNNNNGTC-----ATCGCCACCACTCTTCGCGCCACAGCAAGGGAGGCTTC 809
 QY GCGCGGATGGGTATGCCAGACCAAGTAAAGTGGAGTTGTTGGTACATCTGA 360
 DB GCCGCTCGGCTCGCGCGCTCTCGGCGCGCGCTCGCGCTCANNNNNNNNNNNN 869
 QY CCAGGGGGAATCTGATGATGCGGCAATGGCCATGCCATTTGGACTGGTGGCCATG 420
 DB NNNNNNNNNNNNNNNNNAGTCTCGCGCTGCGCCAGCGCGCTCGACTCGGTCCTCATG 929
 QY GTGGTATTTCTGAGAGGTGGGCGGTCGATGATGTTGAGCGATGCTTTCCAGAAAT 480
 DB GTGCGCATCAGGGAGAGGTCCGCGAGCATGATGTGGACCGACGCTTCCAGGAGACG 989
 QY GACATTTTGGCATCACTTACCGATGTTAGACATCTTATGTGTAGTGTAGTGGCGG 540
 DB CCCATGTCGAGGTCAACCGGCTCATCAACCAAGCAAACTGCTGTCGACGTGAC 1049
 QY GATATGCTCGCATTTGATGAGGCTTTCCATCTTGTAGCAACGGTCGTCGGGCGG 600
 DB GACATNNNNNNCGTGTGTCAGAGGCTTCTCTGCGATCTCTGATCGCGGGGCGG 1109
 QY GTTTGATCGATTTCCCAAGATGTGGGCTTAAAGAAATGTAGTACATTTCCCTCGAC 660
 DB GTGCTGTGCAATCCCAAGACATCC---ACAGAGATGCGGTGCGGCTTGGGA 1165
 QY CCCGCTGAGTTAATCTACCGGGTTATGCCCCAGCTTAAAGTAAATCCCGACAAAT 720
 DB CACGCCCTAGTGTGCTGCTGTTGTGTAATCCCGGGCGCTGCTTATGTGTGGCG 1285
 QY AATGCGGCAAT---TGCAATTTGTGAGCAGCGCAAAATCCTTGTCTTACGTAGGGGG 776
 DB GCTTGAAGAGGTGCTGCTGTTGTGTAATCCCGGGCGCTGCTTATGTGTGGCG 1285
 QY AGGGGCGATCGCCGCGCATGCGCCAGGTGCAAGAAATTCGCGAAAGTTCCAGTT 836
 DB TGG-----CTGCGAGCATTTGGTGAAGATGTGCGCGCTTGTGAGCTGACTGAAT 1339
 QY GCCGCTAACACCACTGATGAGAAATTTGGGCTTTTGCAGAAACCATCCCTTTGCGT 896

DB 1340 CCCGCTCAACACTCTTATGCGCTTGGCAACTTCCCGAGCGACGCCACTGTCTCT 1399
 QY GGGATGTTGGGTATGATGAGCCACCGCTATGCCAACTTCCGCTAGCGAATGTGATT 956
 DB GCGCATGCTGTGTATGATGAGAGGCTTATGCAAAATATGACAGTGTAAAGCGATCT 1459
 QY GTTGAATGCAATGGGGGCGCTTTTGCAGCAGCGGTAACTGCAAACTAGACGAATTTGC 1016
 DB GTTGTCTTCACTTGTGTGCGGTTGAGATGATGCTGTGACAGGAAATGTAGGCTTTTGC 1519
 QY TAGCCGCGCAAGATTAATTCATATGATGATGACCCCGCGAGGTGGGAAACACAGGCG 1076
 DB AGGCAAGGCTAAGATTTGTCAGCTGTGATATGATCCGCTGATTTGGCAAGAACAGCA 1579
 QY TCCCATGTGCCATTTGTTGGGGGATGACCGCATGTTTAAAGACCTTTTGCAGCGGCG 1136
 DB GCCACATGTGTCCATCTGTGAGATGATTAAGCTTGTCTT---GAGAGGCAATGTCTCT 1636
 QY CCGGAAATGATTAATCCACCATCCCATACCAACCGAGCATGTTAAATCCATGTA 1196
 DB TTTTGAAGAGACATCAAGAAAGCTTTGATCTTGTGCTATGGAACGATGAGTTGA 1696
 QY TCATTGGCGGACCGATTACCCCTTCAGAGTSCCCCACTATGAGATTAATTTGCCCCCA 1256
 DB TCAGCAAGAAAGGGAATTTCCCTTGGGTATATAACATCTAATGAGAGATCCAGCCACA 1756
 QY GAGGATGATGACAGAAATTTGTTCGCGCAGGCC---CCGATGCTACTACACGAGATG 1313
 DB ATATGCTATTAAGTTCTTGTATGATGATGACCAAGAGCGCATATGCGCACAGGTGT 1816
 QY GGACAAACCAATATGTGGGCGGCCCATGTTTGA---ACAATGAGCCCGCGCATGGAT 1370
 DB TGGGACAGACCAAGATGTGGGCGGACAGTACTACCTTAAGAGGCGCAAGGAGTGGCT 1876
 QY TTCCAGTGTGCTTGGGTGATGATGAGGCTTTGTGTTACTTCCGCGCATGGAGCCAAAGT 1430
 DB GTCTTACGCTGTCTTGGGCTATGAGATTTGTTNNNNNNNNNNNNNNNNNNNNNT 1936
 QY GGG-AGTGGGGGACAGCGCTCATTTGATGATGAGATGAGATCCAGCTTCCAAATGATCT 1489
 DB GGCACACCAAGTGTACATGTTGTTTACATGATGAGATGATGATGCTTTCCAGAAAGT 1996
 QY TAGGAATCTGGGAACCTTACGCCATGACATCCAGGTAAACTATATTTCTCATTA 1549
 DB TCAGAGAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
 QY CGTTTGGGAGAGGATGTGCGTCAGTGGCAAACTTTCTTACGAAG 1596
 DB CCAGCATCTGGGATGTGTGTCAGTGGAGACAGGTTCTTAAGG 2103

RESULT 2
 CD446001 860 bp mRNA linear EST 03-JUN-2003
 LOCUS E101T0206D04.b Endosperm_4 Zea mays cDNA, mRNA sequence.
 DEFINITION
 ACCESSION CD446001.1 GI:31361644
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 860)
 REFERENCE
 AUTHORS Lai, J., Dey, N., Kim, C.S., Beecraft, P., Larkins, B., Linton, E. and
 Messing, J.
 TITLE Sequencing of the maize endosperm ESTs
 JOURNAL Unpublished (2002)
 COMMENT
 Contact: Lai, Jinseng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801

Mon Jul 26 12:13:50 2004

us-09-893-033-6.1.rst

Page 3

Fax: 732-445-5735
Email: j1a1@waksman.rutgers.edu
Seq primer: 73.

FEATURES
source

Location/Qualifiers
1..860
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="Endosperm 4"
/clone_lib="Endosperm 4"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 10.4%; Score 204.4; DB 14; Length 860;
Best Local Similarity 54.0%; Pred. No. Se-48;
Matches 443; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

236 CCCATGCGCGGATGGGTATGCGCAGACGCAAGTAAAGTGGAGTTTGTTCGGTACAT 355
11 CTTTGGCGGCTCCGCTACGCGCCTCTCCGGGCGCGGTCCGCTCTGCATCCGCACT 70
356 CTGACACGAGGCGCATACCTTGTGACCGGATTCGCAATGCCATTTGGACTCGGTGC 415
71 CCGGCCCCGGCGCCACCACTAGTCTCGGCTGCGGACGCGTGTCTCAATTCCTGTC 130
416 CCATGCTGTGATTACTGAGAGAGTGGCGCTGCCATGATTGTAGCGATCTTCCAG 475
131 CCATGCTGCATCAAGGAGACAGTGGCGCGCATGATTTGGACCGACCTTCCAG 190
476 AAATTGACATTTTGGCATCACTTACCGATTCGTTAAGCATCTCTATGTGTAGTATG 535
191 AGACGCCATGCTGACAGGTCAACCGCTCCATCACCAACACAACTACTGTTCTCGAG 250
536 CGGCGGATATGCGTCGATTTGTTACTGAGGCTTTCATCTTGTCTAGACCGCGTGTCC 595
251 TCGACGACATCCCGCGCGTGGTGGAGAGGCTTTCTCTGCGCTCCCTGTGTGACAG 310
596 GCGCGGTTTATGATGATATCCCAAGATGTGGGCTTGAAGATGTAGTACATTTCCC 655
311 GCGCGGTCTGTGCACATCCCAAGACATCCAGACAGATGGCGGTGCTGTGG 370
656 TCGACCCGCGTACGTTATCTACCGGGTTATCGGCCACGTTAAAGTATCCCGAC 715
371 ACACCCCATAGTGTGCTGGGTACATGGCGCGCTTCCCAAGCCCCCTGCGACTAGT 430
716 AAATTATGCGGATTCGATTTGTTGAGACAGGCGCAAAATCCCTTGTCTACGTAGGG 775
431 TGTCTGACAGAGTGTGCTGTCTGTGTGTAATCCGCGCGCTGTCTTATGTGGGCG 490
776 GAGGGGCGATGCGCGCCCATGCGCCATGCGCGAGTGCAGAAATTGGCGAAAGTTCAGT 835
491 GTGG-----CTGCGAGCATGTGTGAGAGGTTTGCAGCGCTTGTGAGCTGAGAA 544
836 TCGCGGTACCAACCACTGATGGGATTTGGGGCTTTTGAAGAAACCATCCCTTTCCG 895
545 TCCCGGTACCACTACTCTTATGGGCTCGGCACTTCCCAAGAGACCACTGCTGTC 604
896 TGGGTATGTGGGTATGATGGCCACCGCTATGCCAATTGGCCGTACGCGAATGTGATT 955
605 TCGGATCTAGTATGATGAGAGCGGTATGCAAAATTATGCACTGTGATTAAGCCGATC 664
956 TGTATATGACGTGGGGCGCGTTTCAGACACCGGGTACTGCGAACTAGCAATTTG 1015
665 TGTCTCTGCACTTGTGTGGGGTTGATGATCGGTACAGAGAAATTAAGCTTTTG 724
1016 CTAGCCGCGCCAAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1075
725 CAAGCAGGCGCTAAGATTGTGACGCTTGATTTGATTCGCGTGAAGATTGGCAAGAAC 784
1076 CTCCGATGTGCCCATTTGGGGGATGTACGCGCATGTTT 1115

Db 785 AGCCACATGTGTCATCTGTGACAGATTTAAGCTTCTT 824

RESULT 3
CG219527/c 890 bp DNA linear GSS 22-AUG-2003
LOCUS
DEFINITION
O02BK67TV_ZM_0.7.1.5_KB Zea mays genomic clone ZMMEM40753K13,
genomic survey sequence.
ACCESSION
CG219527
VERSION
CG219527.1 GI:34119415
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 890)
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Ronliff, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: O02BK67TV
Contact: Cathy White
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitec@tigr.org
Seq primer: TF
Class: sheared ends.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1..890
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMEM40753K13"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 9.3%; Score 181.6; DB 29; Length 890;
Best Local Similarity 52.4%; Pred. No. 2.5e-41;
Matches 456; Conservative 0; Mismatches 399; Indels 15; Gaps 2;

121 CTTAGCCACGCGCAACCGGGCTTATCTCTGATGATAGCTGAAAGCCATGGGCTC 180
857 CCCACGATCCCGCAAGGCGCGCATCTCTGAGTCCCTGAGCGCTGCGGCGTC 798
181 AAACATATTTTGGCTATCCCGGGGCGCAATTTGCCATCTATATGAATCTGACCG 240
797 CGGAGTCTTCCCTTACCCCGGGCGGTCCATGAGATCCACGAGCACTCAACCGC 738
241 TTGAAGCGCGGGGAAATTGACATATTTTGTGCGGCATGACAGAGAGCTTCCAT 300
737 T-----CCCCGTCATCGCAACACCTCTTCCGCGAGAGAGGGGCGCTTT 687
301 GCGCGGATGGGTATGCGCAGAGCAAGTAAAGTGGAGTTTGTTCGTACATTTGA 360
686 GCGGCTCCCGCTACCGCGGCTCTCTGCGCGCGGTGCGGTGCGATCGCACCTCCGCG 627
361 CCAAGGCGCATACCTTGTGACCGGCAATTTGCCATTTGAGTCTGAGTCCGATG 420
626 CCGGGGCGACCACTTGTCTCGCGCTCGCGGACGCGCTGTGATTCGATTCGATTC 567
421 GTGATATTACTGAGAGGTGGCGCTGCCATGATTGTGAGCGATCTTCCAGAAATT 480
566 GTGCGCATCAAGGAGAGGTGCGGACGACCATATTTGGACCGACCTTCCAGAGAG 507
481 GACATTTTGGATCACTTACCATGTTAAGCACTCTATGTGTAGTGTGAGTGTGGG 540

```

/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCCL1001F04"
/lab_host="XLIBLue MRF"
/clone_lib="CL1"
note="Organ: Pool of sugarcane calli submitted to low

```

111 GGTCGACGAAATTGAACAGTCGGCGCTCTCAGTCTCTGAAATATGACACTCACAGT 52

QY 1241 ATACTATTGCCCCCAGAGGT 1262
Db 51 AAAAGATTAAACCGCAGCGCGT 30

RESULT 5
BH390390/c
LOCUS
DEFINITION BH390390 826 bp DNA linear GSS 11-DEC-2001
AG-ND-153F18, TR ND-TAM Anopheles gambiae genomic clone
AG-ND-153F18, genomic survey sequence.
ACCESSION BH390390
VERSION BH390390.1 GI:17336531
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
1 (bases 1 to 826)
Hon,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J.,
Ren,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B.,
Gardner,M.J. and Collins,F.H.
Construction of a BAC library and generation of BAC end
sequence-tagged connectors for genome sequencing of the African
malaria mosquito Anopheles gambiae
Mol. Genet. Genomics 268 (6), 720-728 (2003)
22542063
12655398
COMMENT Other GSSs: AG-ND-153F18.TF
Contact: Brendan J Loftus
Department of Bukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
This clone is from an A. gambiae BAC library (ND-TAM) provided by
F.H. Collins and sequenced by The Institute for Genomic Research
(TIGR). The BAC library was generated from A. gambiae PBST strain
DNA. All DNA was extracted from newly hatched first instar larvae
to minimize the inclusion of DNA from microorganisms that inhabit
the gut. The DNA is derived from mixed sexes of larvae. The BAC
library was constructed at Texas A&M University BAC Center
University, College Station, Texas 77843-2123, USA using a HindIII
partial digest.
Seq primer: M13 Rev
Class: BAC ends.
Location/Qualifiers
1..826
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PBST"
/db_xref="taxon:7165"
/clone="AG-ND-153F18"
/clone_1fb="ND-TAM"
/note="Vector: pECBAC1; Site_1: HindIII"

ORIGIN
Query Match 9.0%; Score 176.8; DB 28; Length 826;
Best Local Similarity 53.0%; Pred.No.6.1e-40;
Matches 425; Conservative 0; Mismatches 372; Indels 5; Gaps 2;

QY 302 CCGCGATGGTATGCGACAGCCAGGTAAGTGGAGTTTGTTCGGTACATCTGAC 361
Db 826 CAGCGAGGATATGCCCCCTACGATGAAAAAGTGGGTTCGTTTGCACTACCGGTC 767

QY 362 CAGGGGCGACTACTTGTGTGACCGGCATTTGCCAATGCCATTGGACTCGTGCCA--T 419
Db 766 CCGGTGCAACCAATTAGTAAACCGATTGGCTGATGCTCAGATGATTCACACACACTTT 707

QY 420 GGATGGATTACTGGAGAGTGGGGCGGCCCATGATTTGGTAGCATGCTTTCCAGGAAT 479
Db 706 GATGTATTATCCGGGCAAGGTATTGGACATCTTTTGAACAGATGCAATTCAGGAAC 647

QY 480 TGACATTTTGGACATACCTTACCGATCGTTAAGACATCCTATGTGTAGTGGCC 519
Db 646 TGATGTATCAATGTAATGTAATGATGCAAGTACCAAGTGAATTAATCAGTAACTATGCTAC 567

QY 540 GGATATGCTCGCATTTGTTACTAGAGCTTTCATCTTGTAGACACCGGTGTCGGGACC 599
Db 586 ACAGCTCCCAAGCGGTTTACCAAAAAGCATTTCTTATGCGCGTTCCGGAAGACAGAGACC 527

QY 600 GGTTTATGATATTTCCCAAGATGTGGCTTGAAGAAATGATGATCAATTCCTCCGA 659
Db 526 TGTAATATGATATTTCAAAAAATGCCAATTGCAACAATTTGAATA--TCCGGGCTA 470

QY 660 CCGCGTGAAGTTATCTACCGGGTTATGCGCCCAAGGTTAAAGTAAATCCCGACAAAT 719
Db 469 TAGCAATGTGACATTTAGCAATTTGCCCCGAGCCAAATGTAAGAAACGAATATAT 410

QY 720 TAATGCGCATTTGCATTTGTGAGCAGCGCAGAAATCCCTTGCTTACGTAGGGGAGG 779
Db 409 TGAAGAGAGCTGTAATTTATTAATGAAGCCAAAGCCATTTGTATCTTTGACAGAGG 350

QY 780 GGGATGCGCGCCAAATGCCATGCCAGGTGCGAGAAATTTGGGAAAGTTCCAGTTGCC 839
Db 349 AGTCATCTTGGAAGAACAGAAAAAGCTTTCTCAATTTTATGAAAAAGCAATTTTC 290

QY 840 GGTACACCAACCCGTGATGGAAATTTGGGGCTTTTGAAGAAACCATCCCTTCGTGGG 839
Db 289 GGCTGCATGACATATTTATGGGGAGAGTGCATTTCTACTGATCATCCGTTAGAGTTGG 230

QY 900 TATTTGGGTATGATGCGCAACCGCTATGCAACTTTGCCGTAGCGAATGATTTGTT 959
Db 229 TATGTTGGGAATGATGATGATTAACGACCAAAATGTTTAAACAGAAATGATATGTTT 170

QY 960 GATTGCAATGGGGCCCGTTTCAGCAGACCGGGTAACTGCGAAACTGACGAATTTGTAG 1019
Db 169 AATTGCTATAGGAATGATGTTTCCAGCAGCCGTGAACGAGCAATTGCAAAATATGCAA 110

QY 1020 CCGCGCCAAAGTATTCATTCATGATGACATGCCCGCGGAGTGGAAAAACAGGGCTC 1079
Db 109 ACAGGCGACAGGTGATTCATCTGATTAATGATCTCTGCCAAGTGAATTAATAAAC 50

QY 1080 CAAATGCGCATGTTGGGGGAT 1101
Db 49 TACGTTCCGTTTGGGGCGAT 28

RESULT 6
BG595301
LOCUS
DEFINITION BG595301 768 bp mRNA linear EST 07-MAR-2003
EST493979 cSTS Solanum tuberosum cDNA clone CTS9P16 5' sequence,
mRNA sequence.
ACCESSION BG595301
VERSION BG595301.1 GI:13613441
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 768)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chieningo,A.,
Bougri,O., Buell,C.R., Romling,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
TITLE
JOURNAL
COMMENT The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
Location/Qualifiers

source 1..768
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS9P16"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /clone_lib="cSTS"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

Query Match 9.0%; Score 176.6; DB 12; Length 768;
 Best Local Similarity 52.8%; Pred. No. 6.8e-40;
 Matches 408; Conservative 0; Mismatches 359; Indels 6; Gaps 1;

QY 259 ATTGAGCATTTTGTGGCCGATGAAACAGACCTTCCATGGCGGAGATGGTATGCC 318
 Db 2 ATTCGTATGTGCTGCACGTCATGAACAGGGGTGTGTGCTGCAAGGGGTATGCA 61
 QY 319 AGAGCCACAGGTAAAGTGGAGTTTTCGTACATCTGACCCAGGCGACTACTTG 378
 Db 62 CGGGCTACTGGGTTCCCGTGTGTGATCTCATCTGCTCCGGAGGTACGAACTT 121
 QY 379 GTGACCGGATTTGCCAATGCCATTTGGACTCGGTGCCATGTGTGATTACTGGAG 438
 Db 122 GTTAGCGGCTTGGCGAGTCTTGTGTGATGATTCGATTTGCTATTAACGGGTCAA 181
 QY 439 GTGGGCGGTGCGCATTTGTAGAGATCTTCCAGGAAATGACATTTTGGCATACC 498
 Db 182 GTGCCGAGGAGATGATGTGACTGATCCGTTTACGAACTCTATTTGTAGAGTACG 241
 QY 499 TTACCGATCTTAAAGACTCTATGTGTGATGAGTGGCGGAGATATGCTTCGATGTT 558
 Db 242 AGATTCATTAAGAGCATTAATTAATCTGTATGATGATGAGGATATTCCTAGGGTGT 301
 QY 559 ACTGAGGCTTCCATCTTGTAGACACGGGTGTCGGGGCGGTTTGTATGATATTC 618
 Db 302 CGTGAAGCGTTTTCATGAGCAATCGGAGCGGCTGAGCGGTTCTGATTTGATTCCT 361
 QY 619 AAGATGTGGCTTAGAAGATGTGAGTACATTCCTCGACCCCGGTGACGTTAATCTA 678
 Db 362 AAGGATTTACAGAAATGTGATCTAATTTGGATCAGCCATGAGGTGCGTGT 421
 QY 679 CCGGTTATCGCCCAAGGTAAAGTAAATCCCGCAATTAATGGGATTCGATTTG 738
 Db 422 TACATGTAGTGTGCTTAATTTGCTTAATGATGATGCTTTTGGAACTTTGAGCTG 481
 QY 739 TTGAGCAGGCGCAGAAATCCCTTCTACGTAGGGGAGGGCGCATGCGCCAAATGCC 798
 Db 482 ATTTCAAGATGGAAGAGCCCTTTGTATGTGGGTGTGGGTGTA-----CACATCG 535
 QY 799 CATGCCAGGTGACGAAATTTGGCGAAAGTTCCAGTTGCGGTTAACAACACCCCTGATG 858
 Db 536 AGTAGAGAGTGAAGACGATTTGTGAGCTTAACGGGATTCCTGTGGGAGATCTTGTATG 595
 QY 859 GGAATTTGGGCTTTTGAAGAAACCATCCCTTCCGTTGGGTATGTTGGGTATGATGAGC 918
 Db 596 GGTCTTGAATTTTCCATGTGGGATGAGCTTCTCTCAATGTTGGGTAATGATGAGG 655
 QY 919 CACGCTATGCAACTTGTGCGTACGGAATGTGATTTGTTGATTCAGTGGGGGCGCT 978
 Db 656 ACTGTGATGCTATTTATGCGGTGATAGTATGATTTGTTGCTGATTTGGGGTGAAG 715
 QY 979 TTGACGACGCGGATACGCAACTAGAGAAATTTGCTGACGCGCGCCAGT 1031
 Db 716 TTGATGATGAGTTACTGTTAATTTGAAACTTTTGTGACCGGGCGAATTT 768

RESULT 7
 BG599315 775 bp mRNA linear EST 07-MAR-2003
 LOCUS EST504215 cSTS Solanum tuberosum cDNA clone cSTS25A13 5' sequence,
 DEFINITION mRNA sequence.
 ACCESSION BG599315 GI:13619156
 VERSION
 KEYWORDS
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanales; Solanales; Solanum.
 1 (bases 1 to 775)
 van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
 Bougri,O., Bueli,C.R., Romling,C., Tanksey,S. and Baker,B.
 Generations of ESTs from sprouting potato eyes
 Unpublished (2000)
 Contact: Robin Bueli
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: M3F-R.
 location/Qualifiers

FEATURES
 source 1..775
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS25A13"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /clone_lib="cSTS"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

ORIGIN
 Query Match 9.0%; Score 175.6; DB 12; Length 775;
 Best Local Similarity 53.4%; Pred. No. 1.3e-39;
 Matches 395; Conservative 0; Mismatches 339; Indels 6; Gaps 1;

QY 356 CTGACCAAGGGCGACTACTTGTGTGACCGGATTCGCAATGCCATTTGACTGGTGC 415
 Db 2 CTGGTCCGGGAGCTGCAATCTTGTAGTGTCTTGGCGAGTCTTGTGATGATATTC 61
 QY 416 CCATGTGTGATTAATCTGAGAGGTGGCGGTCGATGATGATGATGATGATGATGATG 475
 Db 62 CGATTTGCTATTAACGGGTCAAGTGCAGAGAGATATGATGATGATGATGATGATGATG 121
 QY 476 AATTTGACATTTTGGCATCACTTACCGATGTTAAGCACTCTATGTGTGATGATG 535
 Db 122 AAACCCATTTGTTGAGTAAAGATCTAATTAAGCAATTAATTAATTAATTAATTAATG 181
 QY 536 CCGCGGATATGCTTCGATTTTACTAGAGGCTTCCATCTTGTGATGATGATGATGATG 595
 Db 182 TAGAGGATATTCCTAGGGGTGTGTGAGAGGTTTTCATGCGAAATCGGAGCGCTG 241
 QY 596 GCGCGGTTTATGATATTCCTCAAGATGTGGCTTAAAGAAATGTGATGATATTCGCC 655
 Db 242 GCGCGGTTTATGATATTCCTCAAGATGTGGCTTAAAGAAATGTGATGATATTCGCC 301
 QY 656 TCGACCCCGGTGAGCTTAATCTACCGGCTTATGCGCCCAAGGTTAAAGTAAATCCCGAC 715

Db 302 ATCAGCCATGAGTTCCTGCTGATCATGTCTAGGTACCTTAATTCCTTAATGAGATGC 361
Qy 716 AAATTAATGCGGATTTGCAATTTGTGAGACAGGCGCAAAATCCCTTGCTACGATGAGG 775
Db 362 TTTTGAAACAATTTATGAGGCTGATTTTGAGAGTGAAGAAGCCTGTTTGTATGTGGGTG 421
Qy 776 GAGGGGCGCATGCGCCGCAATGCGCCATGCGCCAGGTGACGAAATTTGGGAAAGTTCCAGT 835
Db 422 GTGGGTGTTTGCAATCAAGTAGG-----AGCTGAGACGATTTGTGAGAGCTTACGGGTG 475
Qy 836 TGCCGCTTAACAACACCCCTGATGAGGATTTGGGGCTTTTGAGCAAAACCATCCCTTTCCG 895
Db 476 TTTCTGTGGCGAGTACTTTGATGAGGTCTTTGAGCTTTTCCAACTGGGAGATGAGCTTTCC 535
Qy 896 TGGATATGTGGATATGATGACGACCGCTATGCGCACTTTGCCGTACGCAATGTGAT 955
Db 536 TTCAATATGTGGATATGATGAGGACCTGTGTATGCTTAATTTATGCTGTGATGTGATGAT 595
Qy 956 TGTGATTTGCACTGGGGGCGCGCTTTCGACGACCGGCTACCTGGCAAACTAGCAATTTG 1015
Db 596 TGTCTCTGCACTTTGGGGGTGAGGTTTGATGATGATGATGATGATGATGATGATGATGAT 655
Qy 1016 CTAGCCGCGCCCAAGTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1075
Db 656 CTAGCCGCGCAAAATTTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 715
Qy 1076 CTCCGATGTGCCCATTTGTG 1095
Db 716 ACCTCATGTTCATTTGTG 735

RESULT 8
CC722291 811 bp DNA linear GSS 19-JUN-2003
DEFINITION OG08B62TV ZM 0.7.1.5_KB Zea mays genomic clone ZMMBMA0697K04,
genomic survey sequence.
ACCESSION CC722291
VERSION CC722291.1 GI:32127067
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 811)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG08B62TH
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: TP
Class: shared ends.
Location/Qualifiers
1..811
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0697K04"
/clone_lib="ZM 0.7.1.5_KB"
/note="Vector: pBSCsk; Site 1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 8.9%; Score 175.2; DB 29; Length 811;
Best Local Similarity 53.5%; Pred. No. 1,8e-39;

Matches 413; Conservative 0; Mismatches 353; Indels 6; Gaps 2;
Qy 275 TGCGCATGACAGAGAGCTTCCATGCGCGGATGGGTATGACAGACAGGTAAG 334
Db 45 TCGGCACAGACAGAGGAGGCTTTCGCGCTCCGCGCTACGCGGCTCTCGGCGCGG 104
Qy 335 TGGAGATTGTTTGGTATCATCTGACACAGGGGCGACTTAATTGTTGACCGGATTTGCA 394
Db 105 TTGGCGCTGACATCGCACCTCCGCGCCGCGCCACCACTAGCTCTGCGCTGCGAG 164
Qy 395 ATGCCCATTTGACCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454
Db 165 AGCGGTGCTGACCTCGCTCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 224
Qy 455 TTGGTACGATGCTTTTCCAGAAATTTGACATTTTGGCATCACTTACCGATGTTAAAC 514
Db 225 TTGGCACCGACGCTTTTCAAGAGACGCCCATGTCAGAGTCAACCGCTCATACCAAGC 284
Qy 515 ACTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 574
Db 285 ACAACTACCTGCTGCTGACGATGACGACATCCCGCGCTGCTGACGAGGCTTTTCC 344
Qy 575 TTGCTAGCACCGGCTGCTCCGCGCGCTTTTATGATGATGATGATGATGATGATGATGATGAT 634
Db 345 TCGCATCTCTGCTGCTCCGCGCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 400
Qy 635 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 694
Db 401 CAGCAGATGCGGCTGCGCGCTGCGGACAGCCCATGATGATGATGATGATGATGATGATGATGAT 460
Qy 695 CGGTTAAAGGTAATCCCGGCAATTAATGCGGATTTGCAATTTGTTGAGACAGGCGAGAA 754
Db 461 CTTCGCAAGCTCCCGCGCTGCGGATTTTGTGAGAGAGTGTGCTGTTGTGATGATGATGAT 520
Qy 755 ATCCCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814
Db 521 CGCGCGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 578
Qy 815 AATTGGCGAAAGGTTCCAGTTCGCGGTAACAACCCGATGATGATGATGATGATGATGATGATGAT 874
Db 579 GCTTTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
Qy 875 ACAGAAACCATCCCTTCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 934
Db 639 CAGCGACGACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Qy 935 TTGCCCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
Db 699 ATCAGAT 758
Qy 995 CTGCAAACTAGACGATTTGCTAGCGCGCCAAATTCATTCATTCATTCATTCATTCATTCATTCAT 1046
Db 759 CAGGAAATTTGAGGCTTTTTCAGGACAGAGCTTAAGTTTGACATTTGAT 810

RESULT 9
CK166070 1121 bp mRNA linear EST 05-DEC-2003
LOCUS CK166070
DEFINITION FGAS050122 Triticum aestivum FGAS: Talc7 Triticum aestivum cDNA,
mRNA sequence.
ACCESSION CK166070
VERSION CK166070.1 GI:38998757
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1121)
Allard,F., Crosby,W.L., Danyluk,J., Endes,F., Frick,M., Gaudet,D.,
Genswein,B., Graf,R., Gulick,P., Hyman,L.D., Laroche,A.,
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
Penniket,C., Roach,J.L. and Sarhan,F.

TITLE
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
JOURNAL
Unpublished (2003)
Contact: Mm L Crosby
COMMENT
Bioinformatics

University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033

Email: fgsa@escc.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. No aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [63,872].
Plate: Talc707 row: J column: 19.

FEATURES

source

1. 1121

/organism="Triticum aestivum"

/mol_type="mRNA"

/db_xref="taxon:4565"

/lab_host="DH5 alpha"

/clone_id="Triticum aestivum FGAS: Talc7"

/note="Organ: Crown; Vector: pGEM-T; SSH (suppression
subtractive hybridization) cDNA library from genotype
C14106 cold hardened at 2 C for 21 days and 49 days
(equal amount of cDNA pooled together before subtraction,
tester) and subtracted against genotype C14106
non-hardened (20 C) (driver). Nitro-pyrole anchored
oligo-dT priming and non-directional cloning."

ORIGIN

Query Match

Best Local Similarity 8.9%; Score 174.6; DB 14; Length 1121;
Matches 423; Conservative 0; Mismatches 349; Indels 9; Gaps 3;

QY 806 AGGTCAGAGATTTGCGGAAAGTTCCAGTCCCGTAAACAACACCTGATGGGAATTG 865
DB 212 AGTTGCGCGCGCTTTGTTGAGCTTACTGGGATTCAGTTACACTACTCTATGGCCTTG 271
QY 866 GGGCTTTGACGAAACCAATCCCTTTGCGGGTATGTTGGTATGCATGCGCACCGCT 925
DB 272 GCAACTTCCCGAGGAGACCACTGCTCTGCGCATGCTGGGATGCATGCACTGTGT 331
QY 926 ATGCAACTTTCGCGTACGAGGATGATTTGTTGATTTGAGTGGGGGCGCGTTTGACG 985
DB 332 ATGCAATTTATGCACTGATTAAGCTTAAGCTTCTGCGATTTGTTGCGGTTGATG 391
QY 986 ACCGGGTAAGTGGCAAACTAGACGAATTTGCTAGCCGCGCAAGTAATTCATTGACA 1045
DB 392 ATCGTGTGACTGGGAAATCGAGGCTTTGCAAGCAGATCCAGATTTGTCACATTGACA 451
QY 1046 TCGACCCGCGGAGTGGGAAAAAAGAGGCTCCCGATGTCCTTGTGGGGGATGTAC 1105
DB 452 TTGACCCAGCTGAAATTTGGCAAGAACAGACCACTGCTTCATTGTCAGATGTA 511
QY 1106 GCCATGTTTGAACACGCTTTTGACGCGGGCCCGGAATGGAATACCCCAACCATCCC 1165
DB 512 AGCTTGTCTT--ACAGGGGTTGAATGATCTAATTAATGGAGCAACACAGAGGTC 568
QY 1166 ATACCAACCCAGGATGTTAATGCACTTATTTGCGGACCGATTAACCCCTCCAGG 1225
DB 569 TGAATTTTGTGTCATGACAAAGAGTTGATCAGCAAGAGAGGATTTCCCTCTGAGAT 628
QY 1226 TGACCCACTATGAGATATATTGCCCCCAGAGAGTATGACAGAAATTTGTCGCGCAG 1285
DB 629 TCAAGACTTTTGGGAGGCACTCCGCGCAATATGATTCAGATTCAGATGAGTGA 688
QY 1286 C---CCCCGATGCTTACTTACACACCGATTTGGAGCAACCAATATTTGGGGCCCAAGT 1342
DB 689 CAAAGGAGGAGGCGCATTTGCCACTGTGTGTGGGCGAGCACCAAGATGTGGGCGGCTCAGT 748

QY 1343 TTTTGA---ACAATGAGCCCCCGCCGATGATTTTCAGAGCTGGCTTGGGTACGATGGGCT 1399
DB 749 ATTACACTTACAGAGGGGCGACGGCAGTGGCTGTCTTCTGCTGTTGGGGCAATGGAT 808
QY 1400 TTGTTTACTTCCCGCATGGAGCCAAAGTGGAGTGGGGGACAGCGCTCATTTGCAT 1459
DB 809 TTGGGTTTACCACTCAGCTGAGCGCTGTGGCCACCAAGTGTTCATGATTTGATGAT 868
QY 1460 CAGTGAAGATGACCACTTCCAAATGATCTTCAGAGACTGGGAACCTTACGCCAGTACA 1519
DB 869 TGATGTATATGTTATTTTCCCTCATGACATTCAGAGATTTGGCGTTGATCCCATTTGACA 928
QY 1520 CATTCAGGTTAAACTATTTATTTTCATATACGTTGGCAGGAGTGTGCTCATGTGCA 1579
DB 929 CCTCCAGTTAAGGTTGATATTTGAACACACAGCATTTGGGAATGTGGGTCCGGGGA 988
QY 1580 A 1580
DB 989 A 989

RESULT 10

CB626704

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2003)

Contact: Rod Ming

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cga cga gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 01 row: C column: 05

Seq primer: gta aac cga cga cga gtc.

Location/Qualifiers

1. 812

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="mRNA"

/cultivar="IR36"

/db_xref="taxon:39946"

/clone="OS11EB01C05"

/tissue_type="leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_id="OS11EB"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (PO6-6-3)"

ORIGIN

Query Match

Best Local Similarity 8.9%; Score 173.2; DB 14; Length 812;
Matches 380; Conservative 0; Mismatches 318; Indels 6; Gaps 1;

QY 415 CCATGATGATTAATCTGAGAGGTGGCGCGTCCATGATTTGGTAGATGCTTTCCAG 474

```

Db      1  CCGATGGTCGCATACCGGCGCAGGTGCCCGCGCAGATCGGACCGAGCCCTTCAG 60
Qy      475 GAATTTGACATTTTGGGATCACTTACCGATCGTTAAGCACTCTATGTGTAGT 534
Db      61  GAGAGCCCATAGTGTAGTGTACCCCGCTCCATCACCAAGCAATTAATCTGTCTGAT 120
Qy      535 GCGCGCATATGCTCGCATTTGTTACTAGAGCTTTCATCTTCTAGCAACGGTCTGCC 594
Db      121  GTGGAGGACATCCCGCGGTACACAGAGCTTCTCTCGGCTCTCGGCGCTCT 180
Qy      595 GGGCGGTTTGTATGATATTCACCAAGGATGGGCTTAGAAGATGTGATTAATGCC 654
Db      181  GCGCGGTGTGTGTGACATCCCAAGACATCCAGCAAGATGGCTGTGCCAGTCTGG 240
Qy      655 CTCGACCCCGGTGACGTTAATCTACCGGGTTATCGCCCGACGTTAAGTATCCCGA 714
Db      241  GACACCTCGATGAATCTACCGGGGTACATGTGACGCTCGCCCAAGCCACCGCGACA 300
Qy      715 CAATTAATGCGGCACTTGCATTTGTGAGCAGGCGCAAAATCCCTGCTTACGTAGG 774
Db      301  TTGCTTGAGCAGGCTCTGCTGTGTGAGGATGACGCGCCCGATTTCTATGTGCT 360
Qy      775 GAGAGGCGGATGCGCCGCAATGCCCATGCGAGGTGACAGAAATTTGCGAAAGTTCC 834
Db      361  GGTGG-----CTGCTTGCACTGTGTGATGATGATGCCCGGTTTGTGAGCTGACC 414
Qy      835 TTGCGGTTAACAACAACCTTGATGGAATTTGGGGCTTTTGAAGAAACCATCCCTT 894
Db      415  ATCCGATTAACAACAACCTGTATGAGGCTCGGCAATTTCCGAGTATGATCCGTTGCC 474
Qy      895 GTGGGTATGTGGGATGATGATGAGCCACCGCATATGCGCACTTTGCGGAGGAAATG 954
Db      475  CTGCGCATGCTTGGGATGATGAGCAGCGTGTACGAAATTTATGCGGTGATAGAGCT 534
Qy      955 TTGTTGATTCAGATGGGGGCCCGTTTTCAGACGACCGGGTACTGTGCAAACTAG 1014
Db      535  CTGTTGCTTGCATTTGGGCGGTGTTGATGATCGTGTGACAGGAAATTTGAGGCTTT 594
Qy      1015 GCTACCCCGCGCAAGTATTCATTCATTCGACCCGCGGAGGTGGGAAAAACAG 1074
Db      595  GCAACGAGGCGCAAAATTTGTCATTCATTCATTCGACGCGAATTTGAAAGAACAG 654
Qy      1075 GCTCCGATGTCGCCATTTGTGGGGATGTAGCGCCATGTTTGA 1118
Db      655  CAACCAATGTCTCAATTTGCGCAGATGTAAAGCTTGCTTACA 698

```

```

RESULT 11
CG970807      897 bp   DNA      linear   GSS 15-DEC-2003
LOCUS      MBENF137F mth2 Medicago truncatula genomic clone 7C2, genomic
DEFINITION survey sequence.
ACCESSION  CG970807
VERSION    CG970807.1 GI:39896586
KEYWORDS
SOURCE
ORGANISM   Medicago truncatula (barrel medic)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
            Medicago.
REFERENCE  1 (bases 1 to 897)
            Town,C.D., Shetty,J., Koo,H. and Felbljum,T.F.
            TITLE      Sequencing of BAC ends from Medicago truncatula
            JOURNAL    Unpublished (2003)
            COMMENT    Other GSSes: MBENF137R
            Contact: Chris Town
            TIGR      9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org

```

```

FEATURES
    source
Seq primer: TGTAAAAACGAGCCAGT
Class: BAC
Location/Qualifiers
    1..897
    /organism="Medicago truncatula"
    /mol_type="genomic DNA"
    /cultur="genotype A17"
    /db_xref="taxon:3880"
    /clone="7C2"
    /clone_1b="mth2"
    /note="Vector: pBeloBAC11, Site 1: HindIII, Site 2:
    HindIII, Cook, D.R. and Kim, D.J., unpublished"

ORIGIN
Query Match      8.7%; Score 170.8; DB 29; Length 897;
Best Local Similarity 51.4%; Pred. No. 3.7e-38;
Matches 456; Conservative 0; Mismatches 417; Indels 15; Gaps 2;

117  TGTTCAGGCAAGGCAAAACGGGGCTTATATCTGATGATAGCTGAACGCCATGG 176
Db    873  TGCCTTACGAACCTCGAAAGGCGCGACATCTCTGATATCTCGAACCTGAGG 814
Qy    177  GGTCAACACATTTTGGCTATCCCGCGGGCAATTTGCCCATCTATGATGACTGTA 236
Db    813  AGTTACAAATGTCTTCGCTTACCCCGCGCGCTTCATAGAGATTCAATCAAGCACT 754
Qy    237  CCGCTTTGAAGCGCGGGGAAATTGACATATTTGGTGGCCATGAAACAGAGCTTC 296
Db    753  GCGCTCCACCGCC-----ATTGTAACGCTCCCTGCTGACGAACAAAGTGAAT 703
Qy    297  CCATGCGGCGAGATGGGTATGCCAGACCAAGGTAAGTGGAGTTTTCGCTACATC 356
Db    702  CTTCGCTCCGAGGATAGCTGCTGCTCTGCTGCTACCTGCTGCTGATGCCCATTC 643
Qy    357  TGAACAAGGGGCACTTACTTGTGACCGGATTCGCAATGCCATTTGACTCGGTGCC 416
Db    642  CCGCCCTGGTGCACATATCTGTCAGTGACACTGTGATGCTTACCTTGACAGTCTTC 583
Qy    417  CATGTGTGATTAAGTGAAGGTGGGCGGTGCGATGATGTGATGATGCTTCCAGGA 476
Db    582  GCTTATGCTATACCGGTGAGGTTCCCGAGAAATGATTTGATGATGCTTTCAAGA 523
Qy    477  AATTGACATTTTGGCATCACTTACCGATCGTTAAGCACTCCATATGTGTAGTGTG 536
Db    522  AACTCGATGTTGAGGTAACTGATCAATCAACCAAGCAATTAATCTGTTGATGAT 463
Qy    537  GCGGATATGCTGCGATTTTACTAGAGCTTTCATCTTGTAGACACGGTGTCCCG 596
Db    462  TGAATATATCAAGATAGTGAAGAGGTTTATTAAGCTTCTAGTGTAGACCTGG 403
Qy    597  GCGGTTTGTATGCAATTCACCAAGATGTGGGCTTAGAAGATGTAGATCTTCCCT 656
Db    402  AACTGTTTGTATGATTAATCTAAGATATTTAGCAACAGGTTTCCCTCCGAATTGG 343
Qy    657  CGAACCCCGGTGACGTTATATACCGGTTATGCCCCACGCTTAAAGTATCCCGACA 716
Db    342  TCAACCTATTAAGTTAACTGGGTATATGAATAGTTTACCMAAGCCCGGATAGGACA 283
Qy    717  AATTAAAGCGGCAATTCATTTGTTGAGCAGGCGCAAAATCCCTGCTCTACCTAGGG 776
Db    282  TTGGAACAGATGTGTAGTGTGTTATTTAGAGTCTAAGAAACCTGTTTGTATTTGG 223
Qy    777  AGGCGCATGCGCGCAATGCCATGCCAGGTGACAGAAATTTGCGAAAGTTCCAGTT 836
Db    222  TGTATGTTT-----GAATGTATGAGGATTAAGAGCGCTTTGTGATTAATCTGAT 169
Qy    837  GCGGTAACGAACACACCTGATTTGGAATTTGGGGCTTTTGAAGAAACCATCCCTT 896
Db    168  TCTGTGTTCTAGTACTTGTATGAGGTTGGCTTATATCTAGTTTGAATGACAAATTC 109
Qy    897  GGGTATGTTGGTATGATGACGACCGCTATGCAACTTTGCGGTGAGGAAATGTATTT 956
Db    108  GCAGATGCTTGGATGATGAGGACAGATTTATGCTAATTTAGCGGTGATTAAGATGAT 49

```

QY 957 GTTGAATTCAGTGGGGGCGCCGTTTCAGACGACCGGTAATGGCAACT 1004
 DB 48 TTTCGTTGGGTTAGGTTTGAATCGGGTTACTGGAGCT 1
 RESULT 12
 CG851756 708 bp DNA linear GSS 18-NOV-2003
 LOCUS ZMMBB0340L06.r ZMMBB Zea mays subsp. mays genomic clone
 DEFINITION ZMMBB0340L06 3', genomic survey sequence.
 ACCESSION CG851756
 VERSION CG851756.1 GI:38378617
 KEYWORDS GSS.
 SOURCE Zea mays subsp. mays (maize)
 ORGANISM Zea mays subsp. mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 708)
 YU, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.,
 and Wang, R.
 TITLE Sequencing of the maize genome
 JOURNAL Unpublished (2003)
 COMMENT Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 440A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR primers
 FORWARD: T7
 BACKWARD: M13r
 Plate: 0340 row: L column: 06
 Seq primer: M13r
 Class: BAC ends.
 FEATURES
 source
 1..708
 location/Qualifiers
 /organism="Zea mays subsp. mays"
 /mol_type="genomic DNA"
 /culti_var="B73"
 /sub_species="mays"
 /db_xref="taxon:4578"
 /clone="ZMMBB0340L06"
 /lab_host="DH10B"
 /clone_lib="ZMMBBb"
 /note="Vector: pRelBAC11, Site_1: HindIII, Site_2:
 HindIII; Zea mays L. ssp. mays"
 ORIGIN
 Query Match 8.7%; Score 170.6; DB 29; Length 708;
 Best Local Similarity 53.6%; Pred. No. 3,7e-38;
 Matches 381; Conservative 0; Mismatches 324; Indels 6; Gaps 1;
 QY 356 CTGACCAAGGGGCACTAATCTGTGACCGGCAATGCCAATTTGACTCGGTGC 415
 DB 705 CGGCCCCCGGCGCACCAACTTGTCTCGGCGTCCGGAGACGGCTCTCGATTCCGTC 646
 QY 416 CCATGAGTGTATTAAGTGAAGGTGGCGGTGCGCATTTGATGAGGAGCTTCCAG 475
 DB 645 CCATGAGTGTATTAAGTGAAGGTGGCGGTGCGCATTTGATGAGGAGCTTCCAG 586
 QY 476 AATTTGACATTTTGGATCACCCTTACCGATCGTTAAGCACTCTGATGTGACGATG 535
 DB 585 AGAGCCCATCGTGAAGGTACCGGCTCATCAACCAAGCAACTGCTGCTCGAGAG 526
 QY 536 CGGCGGATATGCTGCAATTTGATGAGGCTTTCATCTTGGTAGACCGGTCGTC 585
 DB 525 TCGACGATCCCGGCGTGTGAGAGGCTTCTTCTCGCTCTGATGACCGG 466
 QY 596 GGCGGTTTGAATTCGATATTCGCAAGATGTGGGCTTAGAAGATGTGATCATTTCCC 655

DB 465 GCGCGGTGCTTTCACATCCCAAGACATTCACAGACAGATGCGGTGCTGTG 406
 QY 656 TCGACCCCGGTACGCTTATATCAACGGGTATTCGCCCAACGGTTAAAGTATCCCCAC 715
 DB 405 ACAAGCCATGAGTCTGCTGGGTACATTCGCGGCTTCCAAAGCCCCCTGCACTGAGT 346
 QY 716 AATTTAATGCGGCATTTGCAATTTGAGACAGGCCAAGAAATCCCTTGTCTACATGAGG 775
 DB 345 TCTTGAAGAGAGGTGCTGCGTCTTGTGATGATCCCGGCGCTGCTTATATGTGAGCG 286
 QY 776 GAGGGGCGCATGCCCGCCATGCCCATTGCCCGAGTGCAGAGAAATTGGGAAAGTTCCAGT 835
 DB 285 GTGG-----CTGCCAGATCTGTGAGAGATTGCAAGCTTGTGAGCTGACTGAA 232
 QY 836 TCCCGTAAACAACACCTGTATGAGAAATTTGGGCGCTTTTGAAGAAACCATCCCTTTGG 895
 DB 231 TCCCGGTACAACTACTTATTTAGGCGCTCGGAACTTCCCAAGGACAGACCAAGCTGCTC 172
 QY 896 TGGGTATGTGGGTATGATGAGCCACCGCTATGCCACTTTGGCTGCGCAATGTATTT 955
 DB 171 TCGCATGCTAGATGATGATGACAGCGGTATGCAAAATTAATGAGTAAAGCCGATC 112
 QY 956 TGTGATTCAGTGGGGGCGCTTTCAGACGACCGGTAATGCAAACTAGACGAATTTG 1015
 DB 111 TCTTCTTGCATCTGTGTGTGGGCTTGTATGATCGTGTACAGGAAAGTTAGGCTTTG 52
 QY 1016 CTAGCCGGGCCAAAGTAAATTCATTCATTCATGACATGACCGGCGAGGTGGAA 1066
 DB 51 CAAGCAGGGCTMAAAGGTGACGCTGTGATTTGATCCGCTGAGATTTGGCA 1
 RESULT 13
 CNS06TBR 1096 bp DNA linear GSS 05-JUL-2001
 LOCUS T3 end of clone AM0A030D06 of library AM0A from strain CLB 89 of
 DEFINITION Yarrowia lipolytica, genomic survey sequence.
 ACCESSION AL414349
 VERSION AL414349.1 GI:12187440
 KEYWORDS GSS.
 SOURCE Yarrowia lipolytica
 ORGANISM Yarrowia lipolytica
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Dipodascaceae; Yarrowia.
 1 (bases 1 to 1096)
 Soulier, J. L., Aigle, M., Artiguenave, F., Blandin, G.,
 Bolotin-Fukuhara, M., Bon, E., Broctier, P., Casaregola, S.,
 de-Montigny, D., Dujon, B., Durren, P., Lepingle, A., Llorente, B.,
 Malpertuy, A., Neuvéglise, C., Ozier-Lalonde, O., Potier, S.,
 Saurin, W., Tekala, F., Toffano-Nicche, C., Weslowski-Douvet, M.,
 Wincker, P. and Weissenbach, J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 Yeast species for molecular evolution studies
 JOURNAL PERS Lett. 487 (1), 3-12 (2000)
 MEDLINE 11152876
 PUBMED 20584711
 REFERENCE 2 (bases 1 to 1096)
 AUTHORS Casaregola, S., Neuvéglise, C., Lepingle, A., Bon, E., Feyrerol, C.,
 Artiguenave, F., Wincker, P. and Galliardin, C.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
 lipolytica
 JOURNAL PERS Lett. 487 (1), 95-100 (2000)
 MEDLINE 11152892
 PUBMED 20584727
 REFERENCE 3 (bases 1 to 1096)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,

Db 190 CGCCCATCGTCGAGGTACACCCGCTCATACCAAGACAACTACCTGTCCTGACGTCG 249
 Qy 539 CGGATATGGCTCGCATTTTCTACTGAGGCTTTCATCTTGCTAGACACCGGTGTCCTCCGGCC 538
 Db 250 ACGATATCCCGCGCTCGTCAGAGAGGCTTCTTCTCGCATTCCTGTCGCGCCGGGCGC 309
 Qy 539 CGGTTTGTATGATATTCCTCAAGATGCGGCTTAGAAGATGTGATACATTCCTCCCTCG 658
 Db 310 CGGTCTTGTGATCATCCCAAGACAT----CCAGACAGCATGCGGCGTCCGCGCTCG 365
 Qy 659 ACCCGGTGACGTTATATTCACCGGTTATCGCCCAAGGTTAAAGTAAATCCCGACAA 718
 Db 366 GACACGCCCATGATGCTGCTCGGTGATCTCGCGGCTTCCCAAGCTCCCGCATGAA 425
 Qy 719 TTAATGCGGATTTGCAATTTGTTGAGACGCGCAGAAAATCCCTGCTACGTAGGCGGAG 778
 Db 426 TTTCTTGACAGAGTGCTGCTGCTTGTGTGTAATCAAGGCGCTGTTCTTTATGTGGC 485
 Qy 779 GGGCGATGCGCGCAATGCCATGCCAGAGTGACAGAAATTTGCGAAAGTTCCAGTTGC 838
 Db 486 GGTGCTGTGACAGA--TCAGGTAGAGGATGTGCTCCGCTTGTGTGAGTTGACTGGAATCC 543
 Qy 839 CGGTAAACAGACACCTGATGAGAAATGAGGCTTTTGAAGAAACATCCCTTCCGTTG 898
 Db 544 CAGTACACTACTCTTTATGAGGCTTGGCAATCCCGACGACACCACTGTACTGC 603
 Qy 899 GTATGTTGGATATGATGCGCACCGCTATGCCAACTTTCGCTGACGAAATGATTTGT 958
 Db 604 GCATGCTTGTATGATGATGACAGATGATGCAATTTATGAGATGATAGAGCCGATGTGT 663
 Qy 959 TGATTCAGTGGGGGCGCTTTCAGACGACCGGGTAACTGCGCAACTGAGCAATTTGCTA 1018
 Db 664 TGCTTCATTTGTTGTTGGGCTTGTATGATGATGAGACAGGAAATTTGAGGCTTTTGACG 723
 Qy 1019 GCCCGCCCAAGTATTCACATTCGACATTCGACCCGCGG 1057
 Db 724 GCAGAGCTTAAGATTGTGACATGATATGATCTGCTG 762

RESULT 15
 B1935831
 LOCUS EST55720 tomato flower, anthesis Lycopersicon esculentum cDNA
 DEFINITION clone cTOD24K17 5' end, mRNA sequence.
 ACCESSION B1935831
 VERSION B1935831.1 GI:16250303
 KEYWORDS EST.
 ORGANISM Lycopersicon esculentum (tomato)
 SOURCE Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 802)
 van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
 Utechtack,T., Van Aken,S., Romling,C.M., Nierman,W., Fraser,C.M.,
 Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
 Generation of ESTs from tomato flower tissue, anthesis (2001)
 Unpublished (2001)
 COMMENT
 TITLE
 JOURNAL
 CONTACT: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3:
 Location/Qualifiers
 1..802
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultiivar="TA496"
 /db_xref="taxon:4081"
 /clone="cTOD24K17"

/tissue_type="flower"
 /dev_stage="anthesis"
 /clone_lib="tomato flower, anthesis"
 /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; supplier: Cornell University; sequencing: The
 Institute for Genomic Research; flower buds and flowers
 were taken from greenhouse plants (4-8 wks old, TA496).
 They were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

ORIGIN
 Query Match 8.4%; Score 164.6; DB 12; Length 802;
 Best Local Similarity 52.4%; Pred. No. 2.3e-36;
 Matches 389; Conservative 0; Mismatches 349; Indels 5; Gaps 1;
 Qy 259 ATTGACATATTTGTGTGCGGCATGACAAAGAGCTTCCATGCGCGGATGGGATAGCC 318
 Db 65 ATTGATATGTGTACCAAGTATGAGCAAGGTGTGTGTTGCTGACAGAGGTTAGCA 124
 Qy 319 AGAGCCACAGGTAAAGTGGAGTTGTTGCGTATCATCTGACACAGAGGCGCACTACTG 378
 Db 125 CGGCTACTGGGTCCCTGATGTTTGATGCTACTGCTGATCCCGAGGTACAAATCT 184
 Qy 379 GTGACCGGATTTGCCAATGCCATTTGACCTGCGGCCATGCTGTGTATTAAGAGAG 438
 Db 185 GTTATGTGCTTGGCGAATCTTTGTTAGATGATTCGATTTGCTATTACAGTTCAA 244
 Qy 439 GTGGCCCGGCATATTTGATGACGATCTTCCAGAAATTTGACATTTTGGCATCAC 498
 Db 245 GTCCCAAGAGATATTTGATGATGATGCTGTCAGAAAGCCCTATTTGTTAGATAG 304
 Qy 499 TTACCGATGTTAAGACTCCTATGTGTGATGATGATGCGGAGATATGCTGCAATGTT 558
 Db 305 AGATCTATTACCAACATTAATTTCTTTATGATGATGATGATGATGATTTCTTAGGTT 364
 Qy 559 ACTGAGCTTTCATCTGCTAGACACCGGTGTCGCGGCGGCTTTGATGATATTTCC 618
 Db 365 CGTGAGCAATTTTCTTCCGAATCGGACGCGCTGCGCAATTTGATGATGATGAT 424
 Qy 619 AAGATGTGGGCTTAGAAGAAATGTAGTACATTCCTCCGACACCCGCTGACGTTAATCTA 678
 Db 425 AAGATATTTACGAAACAATTTGTGATACCTAATTTGGATGACGCAATGAGTTCCGTGT 484
 Qy 679 CCGGCTATTCGCGCCCAAGGTAAAGTATCCCGCAAAATTAATGGGCGCTTGCATTTG 738
 Db 485 TACATGTCTAGATTACCTAAATTTGCTATGATAAATGCTTTGGAACAATTTGATGCTG 544
 Qy 739 TTGAGCAGGCGCAGAAATCCCTTGTCTACTAGGAGGAGGCGATGCGCCCATATGCC 798
 Db 545 ATTTCGAGTGGAGAAAGCCCTGTTTGTATGTGGGTGTGTGGCTTTCGCAATCA---- 600
 Qy 799 CATGCCAGGTGACAGAAATTTGCGAAAGTTCCAGTTGCCGTTAACAACCACTTGATG 858
 Db 601 -GTGAGGAGCTGAGACGATTTGTGAGCTTACAGATATTCCTGTAGGAGATCTTTATG 659
 Qy 859 GGAATTTGGGCTTTTGAAGAAACATCCCTTTCGAGGATGATGTTGGGTATGATGATGCG 918
 Db 660 GGTCTTGAAGCTTTTCCACTGCGGAGTATGATTTCACTTCAAGATTTGGGTATGATGAT 719
 Qy 919 CACCGTATGCAACTTTGCCGTACGCAATGTATTTGTTGATGATGATGATGATGATGAT 978
 Db 720 ACTGTGTATGCTAATTTATGCTGTGATGATGATGATGATGATGATGATGATGATGAT 779
 Qy 979 TTGACGACGCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
 Db 780 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 802

Search completed: July 25, 2004, 07:28:08
 Job time : 5595.05 secs

CC herbicides. The present sequence is *Synechocystis* sp. strain PCC 6803
XX AHS small subunit gene, 11VN
CC
SQ Sequence 565 BP; 141 A; 138 C; 161 G; 125 T; 0 U; 0 Other;

Query Match 100.0%; Score 565; DB 6; Length 565;
Best Local Similarity 100.0%; Pred. No. 1,2e-165;
Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGAATTTTACCCCAATGCGGCGATGCGCTTTTGGCCCCCAATGAACACACC 60
DB 1 GTGGAATTTTACCCCAATGCGGCGATGCGCTTTTGGCCCCCAATGAACACACC 60
QY 61 CTCTCTGTTTAAAGTGAAGATGAAGCCGAGTGTAAACCCGATTCGCGAATTTGGC 120
DB 61 CTCTCTGTTTAAAGTGAAGATGAAGCCGAGTGTAAACCCGATTCGCGAATTTGGC 120
QY 121 CGCCGTGTTTAACTTGAAGCTTGGCGGTGCGCGGAAACGCGGAGCTTTCC 160
DB 121 CGCCGTGTTTAACTTGAAGCTTGGCGGTGCGCGGAAACGCGGAGCTTTCC 160
QY 181 CGCATCACCATGCTGCTGCGGAGATGAACACCATGCAACATGACCAACCACTC 240
DB 181 CGCATCACCATGCTGCTGCGGAGATGAACACCATGCAACATGACCAACCACTC 240
QY 241 TACCAAGTTGTTAAAGTGAAGATGAAGCAATCAACCACTGCTGCTGGAAGG 300
DB 241 TACCAAGTTGTTAAAGTGAAGATGAAGCAATCAACCACTGCTGCTGGAAGG 300
QY 301 GAATGATGCTGGAAGATGAAGCAATGCGGCAATGCGGCAATGCGGCAATGCTA 360
DB 301 GAATGATGCTGGAAGATGAAGCAATGCGGCAATGCGGCAATGCGGCAATGCTA 360
QY 361 GCCCAAGTATTCGCGGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCCCAAGTATTCGCGGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 TGGGGGACCCGCGGATGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 TGGGGGACCCGCGGATGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 GGTGCTGCAACGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GGTGCTGCAACGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 GAAATCCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565
DB 541 GAAATCCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565

RESULT 2

ABA98820
ID ABA98820 standard; DNA; 1721 BP.

XX ABA98820;

XX 18-JUN-2002 (first entry)

DE Soybean AHS small subunit (clone sdc2c.pk001.b10) coding sequence.

KW AHS; Acetolactate synthase; acetylhydroxy synthase;
KW amino acid biosynthesis; plant; herbicide; AHS small subunit;
KW AHS holoenzyme; crop protection chemical; expressed sequence tag; EST;
ss.

XX Glycine max.

XX Key Location/Qualifiers
FH CDS 88..1433
FT /*tag= a
FT /product= "AHS small subunit"
FT /EC_number= "4.1.3.18"

XX

PN US2001044939-A1.

XX 22-NOV-2001.

XX 08-DEC-2000; 2000US-00732618.

XX 04-JAN-2000; 2000US-0174437P.

XX (ABELL) ABELL L M.

XX (FALCO) FALCO S C.

XX (FAMO) FAMODU O O.

XX Abell LM, Falco SC, Famodu OO;

XX WPI; 2002-113395/15.

XX P-PSDB; ABB08416.

PT New small subunits of plant acetolactate synthase and nucleic acids
PT encoding them, useful in screening for novel crop protection chemicals or
PT potential herbicidal compounds based upon holoenzyme inhibition.

XX Claim 7; Page 37; 57pp; English.

The invention relates to new isolated polypeptides, which comprises an
CC acetolactate synthase (AHS) small subunit. Acetolactate synthase is the
CC first committed step in branched chain amino acid biosynthesis in plants
CC and bacteria. The polypeptides are useful in aiding the discovery of new
CC herbicides that inhibit plant acetolactate synthase activity. The
CC polypeptides are useful for preparing plant AHS holoenzyme, which are
CC useful in screening for potential herbicidal compounds based upon
CC holoenzyme inhibition. The AHS holoenzyme is also useful in screening for
CC novel crop protection chemicals. Particularly, the polypeptides can be
CC used as a target to facilitate design and/or identification of inhibitors
CC of those enzymes that may be useful as herbicides. Thus, the polypeptides
CC could be appropriate for new herbicide discovery and design. The
CC polynucleotides are also useful for producing the polypeptides and for
CC screening for new herbicides. The current sequence represents the soybean
CC AHS small subunit (clone sdc2c.pk001.b10) coding sequence

SQ Sequence 1721 BP; 472 A; 401 C; 391 G; 457 T; 0 U; 0 Other;
Query Match 25.0%; Score 141.2; DB 6; Length 1721;
Best Local Similarity 58.2%; Pred. No. 3e-33;
Matches 285; Conservative 0; Mismatches 203; Indels 2; Gaps 2;

QY 54 AACACCCCTCTCTTTTAAAGTGAAGATGAAGCGGAGTGTAAACCCGATTCGCGACT 113
DB 1005 AACACCGTTATCTCACTTGTGAAGATGCTCTGGAATTTTAAACATTTTACAGAGT 1064
QY 114 ATTGCGCGCGCGGTTTAAATGAGAGCTTGGCGGTGCGGGAACAGGGGGA 173
DB 1065 TTTGCTGAAGAGGCTTAAATGAGAGCTTGAAGTTTACCTGTAAGACAGAGTTGAAG 1124
QY 174 CGTTTCCGATCAGATGCTGCTGCGGCGGATGAAGACCATGCAACAGTACCA 233
DB 1125 ACTTCTCGACTTAACTGCTGCTGCGGCAATGATGATTAATGAAGTTGGTGA 1184
QY 234 GCAACTCTAAGTGTGTTAAAGTGAAGATGAAGCAATGCAACGAACTCCCTGTGT 293
DB 1185 GCACTCTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1244
QY 294 GGAAGGGAATGATGCTGCTGGAAGTGAAGGCGCAATGCGGCAATGCAAGGAGATGAT 353
DB 1245 TGAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1304
QY 354 TGAGCTAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413
DB 1305 TGATATGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1364
QY 414 CGA-ATGCTGCGGAGCGCGGATGAAGTGAAGCAATCTCCAGATGTT-GGCGAAGTTGG 471
DB 1365 GGAGCTTACTGAGATGATGAGCAAGATGATGATGATGATGATGATGATGATGATG 1424

Qy 472 CATTAAGAGTGGCTGCAGACGGGCAAAATTCCTTGGTGGGGAATCCGGCGTCAATAC 531
Db 1425 CATTTGTGAGTGGCGACGACACTGGGCGAATTCCTTAATGCCGAGTCCGGTGTGACTC 1484
Qy 532 GGAATATCTG 541
Db 1485 CAAGTACTTG 1494

RESULT 3

AA199682_33/c
Continuation (34 of 45) of AA199682 from base 3300001 (Mycobacterium tuberculosis strain
WP Sequence split into 45 fragments Locus AA199682 Accession AA199682

WP	Fragment Name	Begin	End
WP	AA199682_00	1	110000
WP	AA199682_01	100001	210000
WP	AA199682_02	200001	310000
WP	AA199682_03	300001	410000
WP	AA199682_04	400001	510000
WP	AA199682_05	500001	610000
WP	AA199682_06	600001	710000
WP	AA199682_07	700001	810000
WP	AA199682_08	800001	910000
WP	AA199682_09	900001	1010000
WP	AA199682_10	1000001	1110000
WP	AA199682_11	1100001	1210000
WP	AA199682_12	1200001	1310000
WP	AA199682_13	1300001	1410000
WP	AA199682_14	1400001	1510000
WP	AA199682_15	1500001	1610000
WP	AA199682_16	1600001	1710000
WP	AA199682_17	1700001	1810000
WP	AA199682_18	1800001	1910000
WP	AA199682_19	1900001	2010000
WP	AA199682_20	2000001	2110000
WP	AA199682_21	2100001	2210000
WP	AA199682_22	2200001	2310000
WP	AA199682_23	2300001	2410000
WP	AA199682_24	2400001	2510000
WP	AA199682_25	2500001	2610000
WP	AA199682_26	2600001	2710000
WP	AA199682_27	2700001	2810000
WP	AA199682_28	2800001	2910000
WP	AA199682_29	2900001	3010000
WP	AA199682_30	3000001	3110000
WP	AA199682_31	3100001	3210000
WP	AA199682_32	3200001	3310000
WP	AA199682_33	3300001	3410000
WP	AA199682_34	3400001	3510000
WP	AA199682_35	3500001	3610000
WP	AA199682_36	3600001	3710000
WP	AA199682_37	3700001	3810000
WP	AA199682_38	3800001	3910000
WP	AA199682_39	3900001	4010000
WP	AA199682_40	4000001	4110000
WP	AA199682_41	4100001	4210000
WP	AA199682_42	4200001	4310000
WP	AA199682_43	4300001	4410000
WP	AA199682_44	4400001	4411529

Query Match 24.5%; Score 138.4; DB 4; Length 110000;
Best Local Similarity 57.1%; Pred. No. 1.4e-31;
Matches 290; Conservative 0; Mismatches 216; Indels 2; Gaps 2;

Qy 22 CACCGCGATGCCCTCTTTCGCCCCCGACGAAACACACCTCTCTGTTTGAAGTGAAGT 81
Db 61145 CACCGAAGCCACGCGCTGATGATGCCCCGAAGACGACAGTGTCTGTTGTCGAAAGAC 61086
Qy 82 GAAGCCGAGTGTCTAACCGCATTCGCGACTATTGGCCGCGGTGTTTAAACATTGAG 141
Db 61085 AAGCCCGGCGTCTGCGCGGAGTGGCGGCGCTGTTTCCCGCGCGGTTTCAACATCGAG 61026

Db 61025 TCCTTGCAGTGGTGGTGCACCGAGTGCAGAGACAGTGCAGTACGATCGCTGCTCC 60966
Qy 202 GGGGATGAGAACACCATCGAACACTGACCAAGCAACTTACAGTTGGTTAACTGTAAT 261
Db 60965 GCCGAGGACACTCCGCTCGAGCAGATCACCAAGCACTCAACAGCTGATCAAGTCATC 60906
Qy 262 AAGTACAGGACATCAACCGAACTCCTGTGTGGAAGGAAATGATGCTGGGGAAGTG 321
Db 60905 AAGTCTCGAGCAGACGACGACGACCTCGTGTCAAGGAAATGGCGTCAACAGTTC 60846

RESULT 4

AA199683_33/c
Continuation (34 of 44) of AA199683 from base 3300001 (Mycobacterium tuberculosis strain
WP Sequence split into 44 fragments Locus AA199683 Accession AA199683

WP	Fragment Name	Begin	End
WP	AA199683_00	1	110000
WP	AA199683_01	100001	210000
WP	AA199683_02	200001	310000
WP	AA199683_03	300001	410000
WP	AA199683_04	400001	510000
WP	AA199683_05	500001	610000
WP	AA199683_06	600001	710000
WP	AA199683_07	700001	810000
WP	AA199683_08	800001	910000
WP	AA199683_09	900001	1010000
WP	AA199683_10	1000001	1110000
WP	AA199683_11	1100001	1210000
WP	AA199683_12	1200001	1310000
WP	AA199683_13	1300001	1410000
WP	AA199683_14	1400001	1510000
WP	AA199683_15	1500001	1610000
WP	AA199683_16	1600001	1710000
WP	AA199683_17	1700001	1810000
WP	AA199683_18	1800001	1910000
WP	AA199683_19	1900001	2010000
WP	AA199683_20	2000001	2110000
WP	AA199683_21	2100001	2210000
WP	AA199683_22	2200001	2310000
WP	AA199683_23	2300001	2410000
WP	AA199683_24	2400001	2510000
WP	AA199683_25	2500001	2610000
WP	AA199683_26	2600001	2710000
WP	AA199683_27	2700001	2810000
WP	AA199683_28	2800001	2910000
WP	AA199683_29	2900001	3010000
WP	AA199683_30	3000001	3110000
WP	AA199683_31	3100001	3210000
WP	AA199683_32	3200001	3310000
WP	AA199683_33	3300001	3410000
WP	AA199683_34	3400001	3510000
WP	AA199683_35	3500001	3610000
WP	AA199683_36	3600001	3710000
WP	AA199683_37	3700001	3810000
WP	AA199683_38	3800001	3910000

Qy 441 GTAGCAATCCTCCAGATGTTGGCCAAGTT-GGCATTAAAGAGGTGCTCGAACGGGCMAA 499
Db 60725 GAGGCCCTGCTGGGGGTGTGGAGCGCTTGGTATTCGGAATCCGCCAATCCGAATG 60666
Qy 500 ATTGCTTGGTGGGGAATCCGGCGTCA 527
Db 60665 GTGTGCTGTCCCGCGGTCCGCGCGCA 60638

WP AA199683_40 400001 411000
WP AA199683_41 410001 421000
WP AA199683_42 420001 431000
WP AA199683_43 430001 4403765

Query Match 24.5% Score 138.4; DB 4; Length 110000;
Best Local Similarity 57.1%; Pred. No. 1.4e-31;
Matches 290; Conservative 0; Mismatches 216; Indels 2; Gaps 2;

QY 22 CACCGCGATCGCTCTTCTTCCCGCCCATGAAACACACCTCTCTGTTTAACTGAAGAT 81
DB 55477 CACCGAGGCGACCGCTGATGAGCCCGAAGACGACAGCTGTGCTGTGTGTAAGAC 55418
QY 82 GAAGCCGAGTGTCTTAAACCCGATTCGCGGACTATTTCGCGCTGTGTTTAACTGAG 141
DB 55417 AAGCCGCGCTGTGCGCGCGGATGCGCGCTGTCTTCCGCGCGCTTCAACATGAG 55358
QY 142 AGCTTGGCGTGTGGGTGCGCGGAAACAGGGGACGTTTCCGCAATCACCATGTGTGCGC 201
DB 55357 TCGTTGGCGGTGTGGGTGCGCGGAAACAGGACGATGACGATGACATGCTGTCTCC 55298
QY 202 GGGGATGAGAACACCATGAAACAACTGACCAAGCACTCTTCAAGTTGTTAACTGAT 261
DB 55297 GCGGAGGACACTCGCTGTGAGCAATGACCAAGCACTCTTCAAGTTGTTAACTGATC 55238
QY 262 AAAGTACAGACATCACCGAACTCCCTGTGTGAAAGGAAATGATGTGTGAAAGTG 321
DB 55237 AAGATCTGTGAGAGAGAGAGAGACATCGCTGTGCAAGGAAATGCGGCTCATCAAGTC 55178
QY 322 AGCGCAATGCGCTCTTACCGAGCGGAAGTATTAAGTACGCGGATTCGCGGCGCC 361
DB 55177 CAAGCGAGCGCGGAGCGCGAGCGCAAGTATGCAAGCGGTAATCTGTTCCGCGCAAC 55118
QY 362 ATTGTGATATCTCCGAAAGACACCGTCCCATGATGATGTTGGGAGACCGG-GGTAAATG 440
DB 55117 GTGATGATGATATCCCGGAGATCATTAACCTCAAGGACACCGGTAACCGCGCAAGTTA 55058
QY 441 GTAGCAATCTCTCAAGATGTTGGGCAAGTT-GGCAATTAAGAGTGTGCTGAAACGCGCAA 499
DB 55057 GAGGCGCTGTGCGGAGTGTGAGACCGGTTGCGTATGCGCAATCGCCCAATCGGAAATG 54998
QY 500 ATTGCTTGTGTGCGGAAATCGCGGCTCA 527
DB 54997 GTGTGCTGTCTCCCGGATCGCGCGGCA 54970

RESULT 5
ID ACA38392 standard; DNA; 504 BP.
XX ACA38392;
AC ACA38392;
XX 19-JUN-2003 (first entry)
DT XX
DE Prokaryotic essential gene #20049.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
OS Mycobacterium bovis.
XX
PN MO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX

PA (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
XX MPI: 2003-029926/02.
DR P-FSDB; ABU34522.
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 26252; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

Sequence 504 BP; 100 A; 151 C; 165 G; 88 T; 0 U; 0 Other;
SQ
Query Match 24.2% Score 137; DB 7; Length 504;
Best Local Similarity 57.7%; Pred. No. 3.5e-32;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;
QY 41 TCCCCCATGAAACACACCTCTCTGTTTAACTGAAGTGAACCGGAGTCTAACCC 100
DB 2 TAGCCCGAAGACGACACGTTGTGCTGTGTGTAAGACAAAGCCGCGTGTGCGC 61
QY 101 GCATTGCCGAGTATTTCGCGCGCGTGTGTTTAACTGAAAGTGTGGGTGCG 160
DB 62 GGTGTGCGCGCTGTCTCCCGCGCGGTTTCAATCAATGATCGTTGGCGGTGCA 121
QY 161 CGAAACAGGGGACGTTTCCGCAATCACCATGTGTGCGGGGATGAAACACATCG 220
DB 122 CCGAGTGAAGGACAGTCAAGATGACATCGTGTCTCCGCGAGACCTCGCTCG 181
QY 221 AACAACTGACAGCACTCTTAAAGTTGTTAACTGAAAGTGTGAGACATCAGC 280
DB 182 AGCAGATCAACAAAGAGCTCAACAGCTGATCAACAGTCAATCAATCTGAGAGAGC 241
QY 281 AAATCTCTGTGTGAAAGGAAATGATGTGTGTAAGTGAAGGCGCAATGCTTAAC 340
DB 242 ACGAGCATCTGCTGTGCAAGGAAATGCGCTCATCAAGTCAAGCCGCGCAAGC 301
QY 341 GAGCGAAAGTATGATGCTAACCGCATGATTCGCGCGCGCAATGTTGATCTCGAAG 400

Db 302 GCAGCCAGTATCGAACCGGTGATCTGTTTCGGCCCAACGTATGACGTATCCCGG 361
Qy 401 ACACCGTACCATGAAATGTTGGGGACCCG-GGTAAATGATGCAATCTCCAGATGT 459
Db 362 AGTATTATACCGTGCAGGCCACCGGTATACCCGGGAATGTAGAGCCCTGCTGGGGTGT 421
Qy 460 TGGCCCAAGTT-GGCATTAAAGAGTGGCTCGAAGCGGCAAAATGCTTTGGTGGGGAAT 518
Db 422 TGGAGCCGTTGCGTATTCGCGAATTCGCCAATCGCGAATGATGTGCTGCCCGCGTCC 481
Qy 519 CCGGCGTCA 527
Db 482 CGCGCGCA 490

RESULT 6

ACA40768
ID ACA40768 standard; DNA; 507 BP.

XX
AC ACA40768;

XX
DT 19-JUN-2003 (first entry)

XX
DE Prokaryotic essential gene #22425.

XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX
KM drug design; gene.

OS
OS Mycobacterium tuberculosis.

XX
PN WO200277183-A2.

XX
PD 03-OCT-2002.

XX
PF 21-MAR-2002; 2002WO-US009107.

XX
PR 21-MAR-2001; 2001US-00815242.

XX
PR 06-SEP-2001; 2001US-00948993.

XX
PR 25-OCT-2001; 2001US-0342923P.

XX
PR 08-FEB-2002; 2002US-00072851.

XX
PR 06-MAR-2002; 2002US-0362699P.

XX
PA (ELIT-) ELITRA PHARM INC.

XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;

XX
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX
DR WPI; 2003-029926/02.

XX
DR P-PSDB; ABU36898.

XX
PS Claim 14; SEQ ID NO 28638; 1766bp; English.

XX
CC The invention relates to an isolated nucleic acid comprising any one of

XX
CC the 6213 antisense sequences given in the specification where expression

XX
CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX
CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX
CC encoding a polypeptide whose expression is inhibited by the antisense

XX
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX
CC polypeptide or its fragment whose expression is inhibited by the

XX
CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX
CC proliferation; (7) identifying a compound that influences the activity of

XX
CC the gene product; or that has an activity against a biological pathway

XX
CC required for proliferation, or that inhibits cellular proliferation; (8)

XX
CC identifying a gene required for cellular proliferation or the biological

XX
CC pathway in which a proliferation-regulated gene or its gene product lies

XX
CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_poc_sequences

XX

SQ Sequence 507 BP; 101 A; 151 C; 166 G; 89 T; 0 U; 0 Other;

Query Match 24.2%; Score 137; DB 7; Length 507;

Best Local Similarity 57.7%; Pred. No. 3, 5e-32;

Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;

Qy 41 TGGCCCCCATGAAACACACCCCTCTCTGTTTAAAGATGAAGCCGAGTGTAAACC 100

Db 2 TGAAGCCGAAGACGACAGTGTGCGTGGTTCGAAGACAAGCCCGCTGCTGAGCC 61

Qy 101 GCATTGCCGACATTTGCCCCCGCGGTTTAAACATGAGAGCTTGCGGGTGGGTCGG 160

Db 62 GGGTGGGGGCGCTTCTCCCGGGGCTTCAACATCAAGTCTGTTGGGGTGGGTCGA 121

Qy 161 CGGAACAGGGGGAGCTTTCCTCCGATCACCATGCTGCTGCGGGAGTGAACACCATCG 220

Db 122 CCGAGTGAAGACAGGTACCGATGACCATGCTGCTCCCGCGAGGACATCCCTCG 181

Qy 221 AACACTGATCAAGCACTTACAGTGTGTTAAGTAAAGTACAGACATCACCG 280

Db 182 AGCAGATCAACCAAGAGCTTCAACAGTCAACAGTCAAGTCTGACAGACAGAG 241

Qy 281 AAATCCCTGCTGGAAGAGGATGATGCTGATGGAAGTGAAGCCCAATGCCCTTAAC 340

Db 242 ACAGACCTGCTGCTACCGGATTTGGCGTATCAAGTCCAGCGAGCGGAGCG 301

Qy 341 GAGCGAAGTGAATGACTGACCCAGATATCCGGCCCGCATTTGATATTCGGAAG 400

Db 302 GCAGCCAGATGATGAAGCCGATGATCTGTTTCGCCCAACGATGATGATGATCCCG 361

Qy 401 AACCCGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 459

Db 362 AGTATTTGACCGTGCAGGCCACCGGTACCGCGGCAAGTTAGAGCCCTGCTCGGAT 421

Qy 460 TGGCCCAAGTT-GGCATTAAAGAGTGGCTCGAAGCGGCAAAATGCTTTGGTGGGGAAT 518

Db 422 TGGAGCCGTTGCGGATTCGGGAATTCGCCCAATCGGAAATGATGATGATGATGAT 481

Qy 519 CCGGCGTCA 527

Db 482 CGCGCGCA 490

RESULT 7

ACA37689
ID ACA37689 standard; DNA; 504 BP.

XX
AC ACA37689;

XX
DT 19-JUN-2003 (first entry)

XX
DE Prokaryotic essential gene #19346.

XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX
KM drug design; gene.

XX
OS Mycobacterium avium.

XX	WM020077183-A2.
PM	
PD	03-OCT-2002.
PB	
PF	21-MAR-2002; 2002MO-US009107.
PR	
PA	(ELIT-) ELITRA PHARM INC.
XX	
PI	Wang L, Zamudio C, Malone C, Hesselbeck R, Ohlsen KL, Zyskind JW,
PL	Wall D, Trizwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPi; 2003-029926/02.
PT	P-PSDB; ABU33819.
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
PT	isolate candidate molecules for rational drug discovery programs.
XX	
P5	Claim 14; SEQ ID NO 25559; 1766bp; English.
XX	
CC	The invention relates to an isolated nucleic acid comprising any one of
CC	the 6'13 antisense sequences given in the specification where expression
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:
CC	(1) a vector comprising a promoter operably linked to the nucleic acid
CC	encoding a polypeptide whose expression is inhibited by the antisense
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC	polypeptide or its fragment whose expression is inhibited by the
CC	antisense nucleic acid; (4) an antibody capable of specifically binding
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC	proliferation; (7) identifying a compound in an operon required for
CC	proliferation; (8) identifying a gene product that influences the activity of
CC	the gene product or that has an activity against a biological pathway
CC	required for proliferation, or that inhibits cellular proliferation; (9)
CC	identifying a gene required for cellular proliferation or the biological
CC	pathway in which a proliferation-required gene or its gene product lies
CC	or a gene on which the test compound that inhibits proliferation of an
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC	product is overexpressed or underexpressed; (12) determining the extent
CC	to which each of the strains is present in a culture or collection of
CC	strains; or (13) identifying the target of a compound that inhibits the
CC	proliferation of an organism. The antisense nucleic acids are useful for
CC	identifying proteins or screening for homologous nucleic acids required
CC	for cellular proliferation to isolate candidate molecules for rational
CC	drug discovery programs, or for screening homologous nucleic acids
CC	required for proliferation in cells other than S. aureus, S. typhimurium,
CC	K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC	prokaryotic essential genes. Note: The sequence data for this patent did
CC	not form part of the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	fip.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 504 BP; 93 A; 146 C; 178 G; 87 T; 0 U; 0 Other;
Query Match	24.0%; Score 135.8; DB 7; Length 504;
Best Local Similarity	56.0%; Pred. No. 8,3e-32;
Matches 277; Conservative	0; Mismatches 217; Indels 1; Gaps 1
OY	41 TGCCCCCATGAAACACACCCTCTGTGTTAGTGAAGATGAAGCGAGTGTAACCC 100
D8	2 TGAGTCGGCAGACGACACGCTGTGTGGCTGCAGGAACAACCGCGGTGTCTGCC 61
OY	101 GCATTCCGGAGCATTTTGCCCGCGGTGTTTTAATTGAGACTTGCGCGTGGCGTGG 160
D8	62 GGATGGCGGCGCTGTTCACCGCGCGGTTTCAATCGAATGCTTGCGGTGGCCA 121

Oy		161	CGAAGCAGGGGACGTTCCTCCGCATCACCATGCTGTGTCGCCGGGGATGAGAACAACATCG	220
Dd		122	CCGAGCAGAAAGGACATGTGCGGATGACGATGTGTGTCCGCCGAGGAGACCCTGCTCG	181
Oy		221	AACACAGTCAACAAGCAACTCTAACAGTTGGTTAAAGTAATAATTAAAGTACAGSACATCACG	280
Dd		182	AGCAGATACCAAGACGCTCAAACAGCTGATCAAGCTATCAAGATGCTTGAGCTGAAG	241
Oy		281	AAATCTCCCTGTGTGAAAAGGAAATTGATGCTGTGTAAGTAGAGCCCAATGCCCTTAAC	340
Dd		242	ACGGCAACTCGGTGTGCGGGGAATTGCGTGTGATCAAGTGTGCGGCCGACGCCGCGCACCC	301
Oy		341	GAGGGAAAGATTTGAGCTACAGCCAGAGTAATCCGGCCCCCGGATGTGTGATATATCCCAAG	400
Dd		302	GCACTGACGGTATCGAGCGGTGAACCTGTTCGGCGCCAAGTATTTGACGTTCCTGCCGG	361
Oy		401	ACAACCTGACCATCGA-ATGTGTGGGGGACCCGGGTAAATGTGATGCAATCCTCCAGATGT	459
Dd		362	AGGGCTGACGATCGAGGCCACCGGTGACCGGGCAGATGAGAGCGCTGTGCGGGTGC	421
Oy		460	TGGCCAAAGTTGGATTTAAAGAGGTGGCTCGAACGGGCAAAATTGCTTTGGTGGCGGAATC	519
Dd		422	TGGAGCCCTTCGGTATCCGTGAGATGTCCAATCGGGAGTGTGTGTGTTCCTCCGCGTTC	481
Oy		520	CGCGCTCAATACGSA 534	
Dd		482	CGCGCGGAATCGGCA 496	
RESULT 8				
ID	ABA98816		standard; DNA; 1813 BP.	
XX	AC	ABA98816;		
DT	18-JUN-2002	(first entry)		
DE	Corn ALS small subunit (clone p0094.cssl172ra) coding sequence.			
XX	XX			
KW	ALS; Acetolactate synthase; acetohydroxy synthase;			
KM	amino acid biosynthesis; plant; herbicide; ALS small subunit;			
XX	ALS holoenzyme; crop protection chemical; ss.			
OS	Zea mays.			
CS				
Key		Location/Qualifiers		
FT	CDS	1..1452		
FT		/tag= a		
FT		/partial		
FT		/product= "ALS small subunit"		
FT		/SC number= "4.1.3.18"		
FT		/note= "no start codon present"		
XX				
US2001044939-A1.				
XX				
PD	22-NOV-2001.			
XX				
PF	08-DEC-2000; 2000US-00732618.			
XX				
PR	04-JAN-2000; 2000US-0174437P.			
XX				
PA	(ABEL/) ABEL L M.			
PA	(PALC/) PALCO S C.			
PA	(FAMO/) FAMODU O O.			
XX				
XI	Abell LM, Falco SC, Famodu OO;			
XX				
DR	WPI; 2002-113395/15.			
XX				
PT	P-PSDB; ABB08412.			
XX				
PT	New small subunits of plant acetylactate synthase and nucleic acids			
XX	encoding them, useful in screening for novel crop protection chemicals or			
XX	potential herbicidal compounds based upon holenzyme inhibition.			
XX				

XX Claim 7, Page 31, 57pp; English.
 CC The invention relates to new isolated polypeptides, which comprises an
 CC acetylactate synthase (ALS) small subunit. Acetylactate synthase is the
 CC first committed step in branched chain amino acid biosynthesis in plants
 CC and bacteria. The polypeptides are useful in aiding the discovery of new
 CC herbicides that inhibit plant acetylactate synthase activity. The
 CC polypeptides are useful for preparing plant ALS holoenzyme, which are
 CC useful in screening for potential herbicidal compounds based upon
 CC holoenzyme inhibition. The ALS holoenzyme is also useful in screening for
 CC novel crop protection chemicals. Particularly, the polypeptides can be
 CC used as a target to facilitate design and/or identification of inhibitors
 CC of those enzymes that may be useful as herbicides. Thus, the polypeptides
 CC could be appropriate for new herbicide discovery and design. The
 CC polynucleotides are also useful for producing the polypeptides and for
 CC screening for new herbicides. The current sequence represents the corn
 CC ALS small subunit (clone p0094.cesst172ra) coding sequence
 CC
 SQ Sequence 1813 BP; 545 A; 370 C; 432 G; 466 T; 0 U; 0 Other;

Query Match 24.0%; Score 135.4; DB 6; Length 1813;
 Best Local Similarity 57.5%; Pred. No. 2e-31;
 Matches 281; Conservative 0; Mismatches 206; Indels 2; Gaps 2;
 DB 54 ACACACCCCTCTGTTTAAAGTGAAGAGCCGAGTGTCAACCCGATTGCGGACT 113
 939 ACATCTCTATCATTTCTTGAATGATGTCCTCGAGTTTCTCATCTTGTAAACGGTGT 998
 114 ATTTGCCCGCGTGTGTTTAACTTGAAGCTTGCGGCTGCGGGAACGCGGGA 173
 999 ATTCTCCAGAGGGGCTCAATATTCAGAGCTTCTGTTGGCCAGCTGAAGAAAG 1058
 174 CGTTCCCGCATCACCAGTGTGTGCGGGGAGTGAAGAACCATTCAGAACACTGACCA 233
 1059 AACTCTGCACTACTGTTGTCTCGAAGCTGATATTCATTCGTAAGCTAGTACA 1118
 234 GCACCTCTACAGTGTGTTAACTTAAAGTACAGACATCAACGAACTCCCTGTGT 293
 1119 TCAACTGTAAAGCTCATGTAGTGTATGAAGTTTACAGATTTCACCACTTACCATTTGC 1178
 234 GGAAGGGGATTTGATGCTGTGAGAGTGAAGGCCAATGCCCTTACCGAGCGGAATGAT 353
 1179 TGCTGAGAGGTTAAAGATCATTAAGTGGCGCAATCTACAGTCCAAAGGATGTCTT 1238
 354 TGAGCTAGCCGAGTATTCGCGGCGGCAATTGTGATATTCGGAAGACCGTCAAC-CA 412
 1239 AGATTTGCTCGATTTTGAAGGACAGAAAGTTGACATATCAGACCAACAATTAACAT 1298
 413 TCGAATGTGGGGGACCCGGGTAAATGATGACAACTCTCCAGATGTTGGCCAAAGT-TGG 471
 1299 ACTGCTCAACCGGAGCATTTGACAGATGTTAGATTGCAAAAGATGTAGAGCGATATGG 1358
 472 CATTTAAGAGTGTGCTCGAAGCGGCAAAATGCTTGTGGGGGGGAATCCGGGCTCAATAC 531
 1359 CATCTGTGAGTTGACGAGACGCGCGGATTTGCTTGTCTCGAGAGTCTGAGATTGACTC 1418
 532 GGAATATCT 540
 1419 CAAGTACCT 1427

RESULT 9
 AC26110
 ID AC26110 standard; DNA; 489 BP.
 AC AC26110;
 XX
 DT 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #7767.
 DE Antisense; ds; prokaryotic essential gene; cell proliferation;
 XX

KW drug design; gene.
 OS Burkholderia mallei.
 PN NC020027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WC-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362599P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DR P-PSDB; AB022240.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 13980; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Sequence 489 BP; 109 A; 140 C; 158 G; 82 T; 0 U; 0 Other;
 Query Match 23.8%; Score 134.6; DB 7; Length 489;
 Best Local Similarity 59.1%; Pred. No. 1.9e-31;
 Matches 230; Conservative 0; Mismatches 159; Indels 0; Gaps 0;
 DB 49 ATGAACACACCCCTCTGTTTAAAGTGAAGAGCCGAGTGTCAACCCGATTGCT 108
 1 ATGAGACACATCATTTCTGCTGCTGAGAAATGAACCGGCGCTGTGCGCGTGTGTC 60
 109 GAACTATTTCGCGCGGTGTTTAACTTGAAGAGCTTGCGGGTGGCGGAGCAAG 168

Db 61 GGTCTGTTCTCGGACGCGGCTACAAACATCGAAACCTTGAACGGTGGCGCGACCGAAAGAC 120
Qy 169 GGGAGCCTTCCCGCATACCATGGTGTGCGCGGGAGTGAAGAACCATCGAACTG 228
Db 121 CAGTCGTGTGGCGCTGACCATGTCTCCATCGGCTCCGACGCGTGAATGAACATC 180
Qy 229 ACCAAGCACTCTCAAGTGTGTTAACTAATTAAGTACAGACATACCGAACTCCC 288
Db 181 ACGAAGCATCTGAACCGCCTATGAGGTGTGAAAGTGTGAGACCTGACCGAGGTGCA 240
Qy 289 TGTGTGAAAAGGAATTTATGCTGTGAAGGTGAGCGCCCAATGCCCTAACCGAGCGAA 348
Db 241 CACATCGAGCGGAGACTATGCTCATCAAGGTGAGGCGAGTGGCGAAGAGCGCGAA 300
Qy 349 GTGATTGAGCTAGCCGACGATGTCGGGCGCGCATTTGATCTCTCGAAGACACCGTC 408
Db 301 ATGAGCGGATGCGCGACATTTCCGCGCGCGCATTCATCGACGTCGAAAGAACTAC 360
Qy 409 ACCATCGAATGTGTGGGGGACCCCGGCTAAA 437
Db 361 ACGATCGAATGTGACGGGGGCGAGCGACAA 389

RESULT 10

ACAA1956
ID ACAA1956 standard; DNA; 492 BP.
XX
XX ACAA1956;
AC ACAA1956;
XX
XX 19-JUN-2003 (first entry)
DT
XX Prokaryotic essential gene #23613.
DE
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
XX Neisseria meningitidis.
OS
XX W0200277183-A2.
PN
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948893.
PR 23-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Tiawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU38086.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids, required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 29826; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 492 BP; 133 A; 118 C; 125 G; 116 T; 0 U; 0 Other;

Query Match 23.7%; Score 133.8; DB 7; Length 492;
Best Local Similarity 60.2%; Pred. No. 3.5e-31;
Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 49 ATGAACACACCCCTCTGTTTAACTGAATGAAGCCGAGGCTAACCCGATGCC 108
Db 1 ATGCGACATATCTTATCTGTTCTGATCGAAACGATAGGTGCGATAGCGGCTGTC 60
Qy 109 GGACTATTGTCGCGCGCGGCTTTTAACTGAGAGCTTGGCGGTGCGCGGAAACAG 168
Db 61 GGTGTTGTTCTGACGCGGATTAACATATGATTTGCGGTGGCGCGACCGAAAGAC 120
Qy 169 GGGAGCCTTCCCGCATACCATGGTGTGCGCGGGAGTGAAGACCATCGAACTG 228
Db 121 AAAACCTGTCCCGATACCATGTTTCAACGCGCGCGAGGTTATGAAACAATC 180
Qy 229 ACCAAGCACTCTACAACTGTTTAACTAATTAAGTACAGACATACCGAACTCCC 288
Db 181 ACCAAGCACTCTAATAATTTGATGATGATTAAGTGTGCTGATTTGAAGAAAGCGCT 240
Qy 289 TGTGTGAAAAGGAATTTATGCTGTGAAGGTGAGCGCCCAATGCCCTAACCGAGCGAA 348
Db 241 TTTGTGAAACGGAATGATGTGTGTAAGTGTGCGCGCGCGCAAGACCGGACGAA 300
Qy 349 GTGATTGAGCTAGCCGACGATGTCGGGCGCGCATTTGATCTCTCGAAGACACCGTC 408
Db 301 TTTTACGCTTGAACCGAATCTACCGGCGCGCATTCATCGACGTAAACCGAGCTAT 360
Qy 409 ACCATCGAA 417
Db 361 ACGATTGAA 369

RESULT 11

AAAB1484/C
ID AAAB1484 standard; DNA; 13732 BP.
XX
XX AAAB1484;
AC
XX
XX 04-DEC-2000 (first entry)
DT
XX
XX N. meningitidis partial DNA sequence gnm_32 SEQ ID NO:32.
DE
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX

OS Neisseria meningitidis.
XX WO20002430-A2.
XX 20-APR-2000.
XX 08-OCT-1999; 99WO-US023573.
XX 09-OCT-1998; 98US-0103794P.
XX 30-APR-1999; 99US-0132068P.
XX (CHIR) CHIRON CORP.
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venier JC,
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,
PI Rapuoli R, Pizza M;
XX WPI, 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
XX in the diagnosis and treatment of N. meningitidis infection and other
XX Neisserial infections, for example, N.gonorrhoea.
XX
XX Claim 7, Page 589-593; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
XX from Neisseria genomic sequences. AA81453 to AA82414 represent
XX specifically claimed Neisseria meningitidis genomic DNA sequences;
XX AA81266 to AA81303 and AA825620 to AA825663 represent Neisseria DNA
XX sequences and their corresponding proteins; AA81254 to AA81259 and
XX AA81304 to AA81321 represent PCR primers used in the isolation of
XX Neisseria meningitidis DNA sequences; and AA81322 to AA81452 represent
XX Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
XX used in the exemplification of the present invention. The nucleic acid
XX sequences, protein sequences, and antibodies against them, can be used in
XX the manufacture of a composition. The composition can be used as a
XX medicament (or in the manufacture of a medicament) for treating,
XX preventing or diagnosing infection due to Neisserial bacteria. For
XX example, some of the identified proteins could be components of vaccines
XX against Meningococcus B; against all serotypes; and/or against all
XX pathogenic Neisseriae. Identification of sequences from the bacterium
XX will also facilitate production of biological probes, particularly
XX organism-specific probes. Attempts to make efficacious Meningococcus B
XX vaccines have failed mainly due to antigen tolerance. Multivalent
XX vaccines have also been tried but none have successfully overcome
XX antigenic variability. The provision of further, complete sequences may
XX provide an opportunity to identify secreted or surface exposed proteins
XX that may be presumed targets for the immune system and which are not
XX antigenically variable or at least more conserved than other more
XX variable regions
XX
XX Sequence 13732 BP; 3488 A; 3771 C; 3357 G; 3116 T; 0 U; 0 Other;
XX
XX Query Match 23.7%; Score 133.8; DB 3; Length 13732;
XX Best Local Similarity 60.2%; Pred. No. 1.5e-30;
XX Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 289 TGTGTGAAGGAATGATGCTGTGATGAGTGGCCCAATGCCCTAACCGAGCGGA 348
DB 5621 TTGTGCGACGCAACTGATGTTGTGTAAGTCCGTGCGCCGCAAGACCGCGACGAA 5562
QY 349 GTGATGAGCTAGCCGAGTATTCGGGCCCGCATTTGTGATATCTCCGAAGACACCTGC 408
DB 5561 TTTTACGCTTGAACCAATCTACCGGGGAGAGATATGACGTAACCGACCGAGCTAT 5502
QY 409 ACCATCGAA 417
DB 5501 ACGATTGAA 5493

RESULT 12
Continuation (2 of 9) of AA81489 from base 100001 (N. meningitidis partial DNA sequenc
WP Sequence Split into 9 fragments LOCUS AA81489 Accession Aa81489
WP Fragment Name Begin End
WP AA81489_0 1 110000
WP AA81489_1 100001 210000
WP AA81489_2 200001 310000
WP AA81489_3 300001 410000
WP AA81489_4 400001 510000
WP AA81489_5 500001 610000
WP AA81489_6 600001 710000
WP AA81489_7 700001 810000
WP AA81489_8 800001 837096

Query Match 23.7%; Score 133.8; DB 3; Length 110000;
Best Local Similarity 60.2%; Pred. No. 3.8e-30;
Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 49 ATGAACACACCTCTCTGTTTAAAGTGAAGACCGGAGTGTAAACCCCATTTGCC 108
DB 102597 ATGGACATATCTTATCTGTTGATTGAAGAAAGATCAAGTCCATGACCCCGTGTG 102538
QY 109 GGAATTTTCCCGCGGTGTTTAAACATTGAGAGCTTGCGGTGGGTGCGGAACAG 168
DB 102537 GGTGTGTTCTCTGACCGCATTAATCAATGATCTTTTGGCGGTGCGCGACGGAAGC 102478
QY 169 GGGGAGCTTCCCGCATCAACATGTGTGTCGCGGGGATGAGAACCAATGAAACTG 228
DB 102477 AAAACCTGTACAGATGACCAATGCTTACCAACGCGGACGAGTGTATGAAATTT 102418
QY 229 ACCAAGCACTTCAAGATTGTTAAAGTAAATTAAGTACAGACATCACCGAACTGCC 288
DB 102417 ACCAAGCACTTCAATTAATGATTGAAGTAAATTAAGTGAATTAAGACCGT 102358
QY 289 TGTGTGAAGGGAATGATGCTGTGTAAGTGAAGGCGCAATCCCTTAACCGAGCGGA 348
DB 102357 TTGTGCAACGCAACTGATGTTGTGTAAGTCCGTGCGCGCAAGACCGGACGAA 102298
QY 349 GTGATGAGCTAGCCGAGTATTCGGGCCCGCATTTGTGATATCTCCGAAGACACCTGC 408
DB 102297 TTTTACGCTTGAACCAATCTACCGGGGAGAGATATGACGTAACCGACCGAGCTAT 102238
QY 409 ACCATCGAA 417
DB 102237 ACGATTGAA 102229

RESULT 13
Continuation (3 of 9) of AA81489 from base 200001 (N. meningitidis partial DNA sequenc
WP Sequence Split into 9 fragments LOCUS AA81489 Accession Aa81489
WP Fragment Name Begin End
WP AA81489_0 1 110000
WP AA81489_1 100001 210000
WP AA81489_2 200001 310000
WP AA81489_3 300001 410000
WP AA81489_4 400001 510000
WP AA81489_5 500001 610000
WP AA81489_6 600001 710000

WP AAA81489.7 700001 810000
WP AAA81489_8 800001 837096

Query Match 23.7%; Score 133.8; DB 3; Length 110000;
Best Local Similarity 60.2%; Pred. No. 3.8e-30;
Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 49 ATGAAACACACCCCTCTGTTTGTAGTGAAGATGAGCGGAGTGTACCCGATGCC 108
DB 2557 ATGCACATATCTATCTGTTCTGATTGAAAACGATCAGTGGATGAGCCGGTGTG 2538

QY 109 GGACTATTGGCCCGCGTGTGTTTAACTTGAAGAGCTTGGCGTGGCGGAAACAG 168
DB 2537 GGTGTGTTCTGTGACGCGGATTAACATTCATCTTTGGCGGTTGGCGGACGAAAGAC 2478

QY 169 GGGGACGTTTCCCGATCACCATGCTGCTGGCGGAGATGAAACACCATCGAACACATG 228
DB 2477 AAAACCTGTACGAGATACCATGCTTACCCAGCGAGCGAGTGTATCGAACAAATT 2418

QY 229 ACCAAGCACTCTACAGATTGTTAACTTAACTTAAAGTACAGACATCACCAGAACTGCC 288
DB 2417 ACCAAGCACTCTACATTAATTGATTGAAGTGAATTAAAGTGTGATTTGAATGAAAAGCGGT 2358

QY 289 TGTGTGAAAGGAAATTGATGCTGTGTAAGGTGAGCGCCAAATGCCCTTACCGAGCGGAA 348
DB 2357 TTTGTGAAACCGAACTGATGTTGTAAGTCCGTCCGCGGCAAAAGCCGCGACGAA 2298

QY 349 GTGATTGAGCTAGCCGAGATATCCGGCCCGCATTTGATATCTCCGAGACACCGTCC 408
DB 2297 TTTTACGCTTGACCGAAATCTACCGGGGAGCATTCATCGATACGATACCGACGCTAT 2238

QY 409 ACCATCGAA 417
DB 2237 ACGATTGAA 2229

RESULT 14
AAAF21611/c
ID AAF21611 standard; DNA; 349980 BP.

XX AAF21611;
XX
XX
XX 13-MAR-2001 (first entry)
XX
XX DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
XX
XX KM Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.
XX
XX OS Neisseria meningitidis.
XX
XX PN WO200066791-A1.
XX
XX PD 09-NOV-2000.
XX
XX PF 08-MAR-2000; 2000WO-US005928.
XX
XX PR 30-APR-1999; 99US-0132068P.
XX 08-OCT-1999; 99WO-US023573.
XX 28-FEB-2000; 2000GB-00004695.
XX
XX PA (CHIR) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX PI Pizsa M, Hickey E, Peterson J, Tertelin H, Venter JC;
PI Maitigani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarsiato V;
PI Rappuoli R, Frazer CW, Grandi G;
XX
XX DR WPI; 2000-647603/62.
XX
XX PT Neisseria meningitidis B full length genome sequence and open reading
frames are used to detect, treat and prevent Neisserial infections.

XX
PS Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used

XX
SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;
Query Match 23.7%; Score 133.8; DB 3; Length 349980;
Best Local Similarity 60.2%; Pred. No. 5.3e-30;
Matches 222; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 49 ATGAAACACACCCCTCTGTTTGTAGTGAAGATGAGCGGAGTGTACCCGATGCC 108
DB 137826 ATGCACATATCTATCTGTTCTGATTGAAAACGATAGATGATGCGTGTGTC 137767

QY 109 GGACTATTGGCCCGCGTGTGTTTAACTTGAAGAGCTTGGCGTGGCGGAAACAG 168
DB 137766 GGTGTGTTCTGTGACGCGGATTAACATTCATCTTTGGCGGTTGGCCGACGAAAC 137707

QY 169 GGGGACGTTTCCCGCATCCATGCTGTGTCGGGGGATGGAACACCATGAAACAATG 228
DB 137706 AAAACCTGTCAAGATGATGCTTACCCAGCGGAGGAGGTATGAAACAAATT 137647

QY 229 ACCAAGCACTCTACAGATTGTTAACTTAAAGTACAGAGCATCAGCAAACTGCC 288
DB 137646 ACCAAGCACTCTACATTAATTGATTGAAGTGAATTAAGTGTGATTTGAATGAAGCGGT 137587

QY 289 TGTGTGAAAGGAAATTGATGCTGTGTAAGGTGAGCGCAATGCCCTTACCGAGCGGAA 348
DB 137586 TTTGTGAAACCGGAATGATGTTGTAAGTCCGTCCGCGGCAAAAGACCGGACGAA 137527

QY 349 GTGATTGAGCTAGCCGAGTATTCGGGCGCGCATTTGATATCTCCGAAAGACACCGTCC 408
DB 137526 TTTTACGCTTGACCGAAATCTACCGGGGAGCATTCAGATCGATCCGACGCTAT 137467

QY 409 ACCATCGAA 417
DB 137466 ACGATTGAA 137458

RESULT 15
ABA9814
ID ABA9814 standard; DNA; 1297 BP.
XX
XX ABA9814;
XX
XX
XX 18-JUN-2002 (first entry)
XX
XX DE Corn ALS small subunit (clone csn3n.pK0112.c11) coding sequence.
XX
XX KM ALS; Acetolactate synthase; acetoaldehyde synthase;
XX amino acid biosynthesis; plant; herbicide; ALS small subunit;
XX ALS holoenzyme; crop protection chemical; ss.

Mon Jul 26 12:13:48 2004

us-09-893-033-17.rng

Page 11

XX Zea mays.

FH	Key	Location/Qualifiers
FT	CDS	
		2. .1054

2.1054

2.1054

```
/*tag=
/part:1a]
```

```
/*tag=
/part:1a]
```

```
/*tag=
/part:1a]
```

```
/product="ALS small subunit"  
/EC number="4.1.3.18"
```

```

/product= "ALS small S
/EC number= "4.1.3.18"

```

```
/note= "no start codon present"
```

PN US2001044939-A1.

PD 22-NOV-2001

PF 08-DEC-2000; 2000US-00732618.

PR 04-JAN-2000; 2000US-0174437P.

PA (ABEL/) ABELL L M.

PA (ABEL/) ABELL L M.

PA (FAMO/) FAMODU O O.

PI Abell LM, Falco SC, Famodu OO;

DR WPI; 2002-113395/15.

DR P-PSDB; ABB08410.

PT New small subunits of plant acetoacetate synthase and nucleic acids
PT encoding them, useful in screening for novel crop protection chemicals or
PT potential herbicidal compounds based upon holotoenzyme inhibition.

PS Claim 7; Page 28; 57pp; English.

The invention relates to new isolated polypeptides, which comprises an acetoacetate synthase (ALS) small subunit. Acetoacetate synthase is the first committed step in branched chain amino acid biosynthesis in plants and bacteria. The polypeptides are useful in aiding the discovery of new herbicides that inhibit plant acetoacetate synthase activity. The polypeptides are useful for preparing plant ALS holoenzyme, which are useful in screening for potential herbicidal compounds based upon holoenzyme inhibition. The ALS holoenzyme is also useful in screening for novel crop protection chemicals. Particularly, the polypeptides can be used as a target to facilitate design and/or identification of inhibitors of those enzymes that may be useful as herbicides. Thus, the polypeptides could be appropriate for new herbicide discovery and design. The polynucleotides are also useful for producing the polypeptides and for screening for new herbicides. The current sequence represents the corn ALS small subunit (clone ccm3n.pK0012.c1) coding sequence

SQ Sequence 1297 BP; 356 A; 263 C; 303 G; 375 T; 0 U; 0 Other;

Query Match	Score	DB	Length
23.5%	132.8	6	1297

Matches	279;	Conservative	0;	Mismatches	207;	Indels	2;	Gaps	2;
---------	------	--------------	----	------------	------	--------	----	------	----

QY 55 CACACCCCTCTGTTTGTGAGATGAAGCCGGAGTGCTAACCCGCATTGCCGACTA 114

Db 542 CATACCTCTCCATCCTGTGAATGATTGTCCTGGTGTCTCAACATTGTAACAGGAGTC 601

115 TTGCCCCGCGTGGTTTACATTGAGAGCTTGGCGGTGGGTCGGCGGAACAGGGGAC 174

Db 602 TTGCTCGCAGGGCTACATATACAGAGCCTTGCTGTTGGCTCAGCTGAGAAGGAGGC 661

175 GTTCCGCATCACCATGGTGGTCCCGGATGAGAACACCATCGAACAACTGACCAAG 234

Db 662 ATTCACGATTACACAGTTGTTCTGCTACTGTGAATCCATTGGGAAGTTAGTTCAG 721

235 CAACTCTACAGTTGGTTAACGTAATTAAGTACAGGACATCACCGAACTCCCTGTGTG 294

Db 722 CAGCTTACAGCTTATTGATGTGCATGAAGTTCATGACATTACCCACTCACCCTTTTGCT 781

295 GAAAGGAAITGATGCTGTGAGGTTGAGCCCAATGCCCCCTAACCGAGCGGAAGTGATT 354

Db

782 GAAAGGAGACTGATGCTTATTAGGTTTCGTAAACACTGCTGCTGGAGGGAATTCTA 841

355 GAGCTAGCCCAAGTATTCGGGCCCGCATTTGTGATATCTCCGAGACACCGTCA-CCAT 413

Db 842 GATATTGCTGAATCTTCCGAGCAAAACCTATTGACGTTTCTGACCATACGTAACCTT 901

414 CGAATGCTGGGGACCCGGGTAAATGCTAGCAATCCTCCAGATGTT-GGCCAAGTTGGC 472

Db 902 CAGCTTACTGGAGATCTTGACAAGATGGTTGCACTACAAGGTTATTAGAGCCATATGGC 961

473 ATTAAGAGGTGGCTCGAACGGCAAAATGCTTTGGTCCGGGAATCCGGCTCAATACG 532

Db 962 ATCTGCGAGGTCGCCAGACTGACGAGTGGCACTGTCCGTGAATCGAAGTGCCTCC 1021

Qy 533 GAATATCT 540

Db 1022 AAGTACCT 1029

Search completed: July 25, 2004, 05:28:02
Job time : 241.282 secs

Job time : 241.282 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 05:10:12 ; Search time 44.0987 Seconds

(without alignments)
7110.132 Million cell updates/sec

Title: US-09-893-033-17
Perfect score: 565
Sequence: 1 gtggaatttaccacatg.....cccggaattccaagttag 565

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	138.4	24.5	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
C 2	138.4	24.5	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
C 3	112.8	20.0	540	4 US-09-252-991A-7249	Sequence 7249, Ap
C 4	112.8	20.0	1722	4 US-09-252-991A-7450	Sequence 7450, Ap
C 5	112.8	18.7	1673	4 US-09-426-568A-1	Sequence 1, Appl1
C 6	105.8	18.5	525	4 US-09-489-039A-5192	Sequence 5192, Ap
C 7	104.6	18.5	507	4 US-09-328-352-3858	Sequence 3858, Ap
C 8	100.2	17.7	1830121	4 US-09-557-884-1	Sequence 1, Appl1
C 9	100.2	17.7	1830121	4 US-09-643-990A-1	Sequence 1, Appl1
C 10	90.2	16.0	546	4 US-09-540-236-676	Sequence 676, App
C 11	90.2	16.0	96109	4 US-09-596-002-35	Sequence 35, Appl
C 12	78	13.8	640681	4 US-09-790-986-1	Sequence 1, Appl1
C 13	74.8	13.2	6211	4 US-08-916-421B-1	Sequence 8, Appl1
C 14	70	12.4	1664976	4 US-08-916-421B-1	Sequence 1, Appl1
C 15	61.6	10.9	4895	4 US-09-426-568A-3	Sequence 3, Appl1
C 16	55.2	9.8	477	1 US-08-403-866-13	Sequence 13, Appl
C 17	55.2	9.8	2231	1 US-08-403-866-14	Sequence 14, Appl
C 18	55.2	9.8	12720	1 US-08-403-866-11	Sequence 11, Appl
C 19	55.2	9.8	12720	1 US-08-403-866-11	Sequence 11, Appl
C 20	34.6	6.1	897	2 US-08-232-463-14	Sequence 14, Appl
C 21	34.4	6.1	1356	4 US-09-252-991A-13894	Sequence 1, Appl1
C 22	34.4	6.1	3795	4 US-09-252-991A-13448	Sequence 13448, A
C 23	34.4	6.1	3795	4 US-09-252-991A-13448	Sequence 13448, A
C 24	33.8	6.0	291	4 US-09-489-039A-3462	Sequence 13693, A
C 25	33.2	5.9	1539	4 US-08-277-031B-6	Sequence 3462, Ap
C 26	33.2	5.9	1539	4 US-08-277-031B-6	Sequence 5, Appl1
C 27	33.2	5.9	1539	4 US-08-277-031B-6	Sequence 6, Appl1
C 27	33.2	5.9	1539	4 US-08-277-031B-7	Sequence 7, Appl1

28	33.2	5.9	2565	4 US-09-023-655-1052	Sequence 1052, Ap
29	33.3	5.8	288	4 US-09-543-681A-1575	Sequence 1575, Ap
30	32.8	5.8	885	4 US-09-252-991A-7525	Sequence 7525, Ap
31	32.8	5.8	969	4 US-09-252-991A-7592	Sequence 7592, Ap
C 32	32.8	5.8	1101	4 US-09-252-991A-7592	Sequence 7592, Ap
C 33	32.4	5.7	633	4 US-09-252-991A-7682	Sequence 7682, Ap
C 34	32.4	5.7	2451	4 US-09-252-991A-7534	Sequence 7534, Ap
C 35	32.4	5.7	2661	4 US-09-252-991A-7807	Sequence 7807, Ap
C 36	31.4	5.6	2294	4 US-09-489-039A-726	Sequence 736, App
C 37	31.4	5.6	618	4 US-09-252-991A-8221	Sequence 8221, Ap
C 38	31.4	5.6	654	4 US-09-252-991A-8152	Sequence 8152, Ap
C 39	31	5.5	1080	4 US-09-125-642C-3	Sequence 3, Appl1
C 40	31	5.5	4332	3 US-08-728-603-14	Sequence 14, Appl
C 41	31	5.5	5515	4 US-09-125-642C-8	Sequence 8, Appl1
C 42	31	5.5	5519	4 US-09-125-642C-12	Sequence 12, Appl
C 43	31	5.5	5904	1 US-08-309-512-1	Sequence 1, Appl1
C 44	31	5.5	5904	1 US-08-309-512-1	Sequence 1, Appl1
C 45	31	5.5	32207	2 US-08-770-379-20	Sequence 20, Appl1

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103, 840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Query Match 24.5%; Score 138.4; DB 3; Length 4403765;
Best Local Similarity 57.1%; Pred. No. 2.2e-33;
Matches 290; Conservative 0; Mismatches 216; Indels 2; Gaps 2;
C 22 CACCGGAGTGGCGCTTGTGGCCCCCAGTAACACACCTCTCTGTAGTTGAAGAT 81
Db 3355477 CACCGGAGTGGCGCTTGTGGCCCCCAGTAACACACCTCTCTGTAGTTGAAGAT 81
C 82 GAAAGCGGAGTGGCGCTTGTGGCCCCCAGTAACACACCTCTCTGTAGTTGAAGAT 141
Db 3355417 GAAAGCGGAGTGGCGCTTGTGGCCCCCAGTAACACACCTCTCTGTAGTTGAAGAT 141
C 142 AGCTGGCGTGGCGTGGCGGGAAGAGGGGAGTGTCCCGCATCCATGCGTGGCGG 201
Db 3355417 AGCTGGCGTGGCGTGGCGGGAAGAGGGGAGTGTCCCGCATCCATGCGTGGCGG 201
C 202 GGGAGTGAAGACACCATGCAACACTGACCAAGCAACTTCAAGTTGTTAACTGTAAT 261
Db 3355297 GGGAGTGAAGACACCATGCAACACTGACCAAGCAACTTCAAGTTGTTAACTGTAAT 261
C 262 AAGTGAAGACACCATGCAACACTGACCAAGCAACTTCAAGTTGTTAACTGTAAT 321
Db 3355237 AAGTGAAGACACCATGCAACACTGACCAAGCAACTTCAAGTTGTTAACTGTAAT 321

```

Query March 20.0%; Score 112.8; DB 4; Length 540;
Best Local Similarity 55.8; Pred. 4,9e-27;
Matches 219; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 48 CATTGAACACACCCCTCTCTGTTTACTTGAAGATGAAGCCGAGTCTTAACCCGCATTGC 107
DB 48 CATGCGCACATCATTTTCCCTGCTGCTGGAAGAACGACCAGGCGCATTTGTCGCGGTGT 107
QY 108 CGGACTATTTGCCCGCCGCGGTGTTTAACATTTGAGAGCTTGAGCGGTGGAGTCGGCGGAACA 167
DB 108 CGGCGTGTCTCCCAACGCAACTAACCATTCGAAAGAGCCTGACCGTGGCGCGACCGAGGA 167
QY 168 GGGGAGCGTTTCCCGCATCAACATGGTGGTCCGGGGGATGAGAAACCATCGAACACT 227
DB 168 CCCGACCCCTGTGGGTGTGACGCTGACCAACCGTGGGCAAGATGAGTGTATCGAGCGAGAT 227
QY 228 GACCAAGCAACTCTACAAGTGGTTAACGTATTTAAAGTACAGACATCAACCGAACTCC 287
DB 228 CACCAAGAACCTCAACAAAGCTGATCGAAGTGGTCAACCTGATGATCTGCGAAGAACGC 287
QY 288 CTGTGTGGAAGGGAATTGATGCTGGTGAAGGTGAGGCGCAATGCCCTTAACGAGCGGA 347
DB 288 CCATATGAGCGCGAGGTGATGCTGGTGAAGGTCAAGGCCACGCGGCCCAAGCGCGCGGA 347
QY 348 AGTGATTAAGCTAGCCCAAGGTATTCGGGCCCCGACATTGTGGATATCTTCGGAAGACCCGT 407
DB 348 GGTCAAGCGCACCAACCATATCTTCCTCGTGGGCAAGATGTGACGTCACCAAGTAGCGTCTA 407
QY 408 CACCATGATGTGGTGGGACCCGGGTAAATAGTGA 443
DB 408 TACCGTGCAACTGGCGGGTACCAAGCAAAACTGGA 443

RESULT 4
US-09-252-991A-7450/c
; Sequence 7450, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A

```

CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7450
LENGTH: 1722
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7450

Query Match 20.0%; Score 112.8; DB 4; Length 1722;
Best Local Similarity 55.3%; Pred. No. 8.9e-27;
Matches 219; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 48 CATGAACACACCCCTCTCTGTTTAAAGATGAAGCCCGAGTGAACCCCGCATTC 107
DB 1654 CATGGAACACATTCCTCTGCTGCTGAAGAAAGAGCCAGCCGATTCCTCCGCGGT 1595
QY 108 CGGACTATTGCCCCCGCTGTTTAAATTGAGAGCTTGGCGGTGGGGTGGCGAACA 167
DB 1594 CGGCGTGTCTCCCAACCACTACACATCGAAAGCTGACCGTGGCGCCGACGAGGA 1535
QY 168 GGGGGACGTTTCCGCTACCATGTGTGGTGGGGGATGAGAACACATCGAACACT 227
DB 1534 CCGGACCTGTGCGCTGAGGCTGACCAACCGTGGGACGATGAGTGCAGCAT 1475
QY 228 GACCAAGCACTTCAAGTGTGTTAACGTAATTAAGTACAGACATCACCGAACTCC 287
DB 1474 CACCAAGAACTCAACAGCTGATGAGTGTGCTGATGCTGTGGAAGAGGC 1415
QY 288 CTGTGTGAAGAGGAAATTGATCGTGTGAAGTGAAGCCCATGCCCCCTAACGAGGGA 347
DB 1414 CCATTCGAGCGCGCTGATGCTGTGAAGTCAAGGCGCGCCGACGCGCGCGA 1355
QY 348 AGTGATTGAGTACCCCAAGTATTCGGGCGCCGCAATTGTGATATCTCCGAACACCGT 407
DB 1354 GGTCAAGGCGACACCGATTCCTCGTGGGAGATCTGCGACGTCACCAATGAGTCTA 1295
QY 408 CACCATCGATGTGGGGGACCCCGGTAAATGTA 443
DB 1294 TACCGTGAACCTGGCGGGTACACAGGACAACTGGA 1259

RESULT 5

US-09-426-568A-1
Sequence 1, Application US/09426568A
Patent No. 6348643
GENERAL INFORMATION:
APPLICANT: Kakefuda, Genichi
APPLICANT: Costello, Colleen
APPLICANT: Sun, Ming
APPLICANT: Hu, Weiming
TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance
TITLE OF INVENTION: In Plants
FILE REFERENCE: 008103/195497
CURRENT APPLICATION NUMBER: US/09/426,568A
CURRENT FILING DATE: 1999-10-22
PRIOR APPLICATION NUMBER: 60/106,239
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 1673
TYPE: DNA
ORGANISM: Arabidopsis sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (42)..(1514)
OTHER INFORMATION: Mature Peptide
US-09-426-568A-1

Query Match 19.8%; Score 112; DB 4; Length 1673;
Best Local Similarity 54.5%; Pred. No. 1.6e-26;
Matches 266; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

QY 55 CACACCCCTCTGTTTAAAGTGAAGCGAGTCTAACCCGATTCGGGACTA 114
DB 999 CATCTCTATCATGCTTTTAATGATATTCAGAGTCTTAATATTTGACTGTGT 1058
QY 115 TTGCCCCCGGTGTTTAAATTGAGCTTGGCGGTGGGGAACGGGAGC 174
DB 1059 TTGCTCGAAGGGATACATATCCAGACTTGGCGGTGAGCATGCTGAACCAAGGCG 1118
QY 175 GTTCCCGGATACCATGTGGTGGCGGGGATGAGAACACATCGAACACCAAG 234
DB 1119 ATTTCAGCATTCACACAGTATACCTGCAACAGATGAATCGGTACGAATTTGTGAG 1178
QY 235 CAATCTACAGTGTGTTAACGTAATTAAGTACAGACATCACCGAACTCCGTGTG 294
DB 1179 CAATTTACAACTGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1238
QY 295 GAAAGGAATTATGCTGTGTAAGTGAAGCGCCATGCCCCCTAACCGAGGAGTAT 354
DB 1239 GAAAGGAATGATGTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1298
QY 355 GACTAGCCGAGTATTCGGGCGCCGATTTGTGATATCTCGAAGACACCGTACAT 413
DB 1299 GACTTGTGATATTTTTCAGGCTTAAGCTGTGATGATGATGATGATGATGATGATGAT 1358
QY 414 CGATGTGGGGGACCCCGGTAATGTGATGATGATGATGATGATGATGATGATGATGAT 472
DB 1359 CAGCTTACGCGGATATCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1418
QY 473 ATTAAGAGTGTGCTGAACCGGCAAAATGCTGTGTGGGGAATCCGCGTCAATAG 532
DB 1419 ATTTGAGGTTTCAAAACCGGTGCTGTGATGATGATGATGATGATGATGATGATGAT 1478
QY 533 GAATATCT 540
DB 1479 AAGTACT 1486

RESULT 6

US-09-489-039A-5192
Sequence 5192, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1998-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 5192
LENGTH: 525
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5192

Query Match 18.7%; Score 105.8; DB 4; Length 525;
Best Local Similarity 54.5%; Pred. No. 9.4e-25;
Matches 212; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 49 ATGAACACACCCCTCTCTGTTTAAAGTGAAGCCGAGTGTAAACCCGATTCGC 108
DB 34 ATGGCCGGAATATTCGATTAATGATGAGAAAGAAATCGGATGATTCGCCGCTTATC 93
QY 109 GACTATTTGCCCCCGGTGTTTAAATTGAGAGCTTGGCGGTGGGGTGGGGAACAG 168
DB 94 GGCTCTTTTCCAGGCTGCTAATATGAAAGCTGCTGCTGCGCGGACGAGCAT 153

QY 169 GGGGACGTTTCCCGCATCAGATGTTGCGGGGGATGAGAACACCATGAACTG 228
DB 154 CCCACTCTCCAGAAATGACATTCAGACCGTGGGCGCATGAAAAGGATGACATATT 213
QY 229 ACCAAGCACTTCAAACTGTTTAACTTAATTAAGTACAGGACATCCGAACTCC 288
DB 214 GAAAAGCAACTGCAAACTGTTGATCTCTGCGGGGACAGGCTGGGGGAGGCT 273
QY 289 TGTGTGAAAGGAAATGATGCTGTTGAGAGTGAAGCCCAATGCCCCCTAACCGAGGAA 348
DB 274 CATGTGAAACGGAAATATCTGTGTGAGGTTGACGGCAGCGCTATGCGCGTGAAGAA 333
QY 349 GTGATTGAGCTAGCCGAGATATCCGGGCGGCAATGTTGATATCTCGAAGACACGCTC 408
DB 334 GTGAAGCCGCAACCCGAATCTCTCCGCGAGATATGACGATGAGCGCTTCTATCTAT 393
QY 409 ACCATGATGATGTTGGGGGAGCCGGGTAAA 437
DB 394 ACCGTCAGCTGCGCGGACCAAGCATATA 422

RESULT 7
US-09-328-352-3858
; Sequence 3858, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3858
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3858

Query Match 18.5%; Score 104.6; DB 4; Length 507;
Best Local Similarity 54.2%; Pred. No. 2.3e-24;
Matches 212; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

QY 49 ATGAAACACACCCCTCTGTTTATGTAAGATGAAGCGGAGTGTACCCGATGCGC 108
DB 16 ATGAGACACATTTATTTCTGTTACTGTTGAAAACGAGCTGTGCGTTTCTCTGTTG 75
QY 109 GCACTATTGCGCGCGTGTGTTTAACTTGAAGCTTGGCGGTGGCGGAGCAG 168
DB 76 GCGTATTCACTCAACGTAACATCAACATTTGAACTTATGTTGACCAACCGAAGAT 135
QY 169 GGGGACGTTTCCCGCATCAGATGTTGCGGGGATGAGAACACCATGAACTG 228
DB 136 CCGACATTATGCGCTTGAAGCTTAACTTATGCGATGACCTTAAGATGAACTAATT 195
QY 229 ACCAAGCACTTCAAACTGTTTAACTTAATTAAGTACAGGACATCCGAACTCC 288
DB 196 ACGAAACAGCTCAACAGCTTGTGAGTGTGTAAGATGATTAATCAGAAAGTTCA 255
QY 289 TGTGTGAAAGGAAATGATGCTGTTGAGAGTGAAGCCCAATGCCCCCTAACCGAGGAA 348
DB 256 CACATGAGCTGAACTATCTATCAAGTAAACATTAAGTGTGTTTCAACGCTGAA 315
QY 349 GTGATTGAGCTAGCCGAGATATCCGGGCGGCAATGTTGATATCTCGAAGACACGCTC 408
DB 316 ATTAAGCAGCTGCGGAAATTTTCCGTGCGCAATTTGAGCGTGAACACCAAGACCTAT 375
QY 409 ACCATGATGATGTTGGGGGAGCCGGGTAAA 437
DB 376 ACCATTCATATGACAGGAGCAACTGAAAAAT 406

RESULT 8

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PA186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 17.7%; Score 100.2; DB 4; Length 1830121;
Best Local Similarity 54.1%; Pred. No. 4.5e-21;
Matches 204; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 62 TCTCTGTTTATGTAAGATGAAGCGGAGTGTACCCGATGCGGACTATTGCCC 121
DB 1651189 TATCTGTTTACTCGAAATGATCGGAGCATTTGCGGTGTTTATTTCTCTC 1651130
QY 122 GCGGTGTTTAACTTGAAGCTTGGCGGTGGGCTGGCGGAAACAGGGGACGTTCCC 181
DB 1651129 AAGTGCTTTAAATTAAGTTGAAGGTTGACCAACGAGATGATCAACGCTTCTC 1651070
QY 182 GCATACATGTTGTTGCGGGGATGAGAACACCATGAACTGAACCAAGCACTCT 241
DB 1651069 GATACACATTTGAACCTGTGGGCGATGCTCAGGATTTGAGCAATGAAACACTTC 1651010
QY 242 ACAAGTTGTTAAGTAAATTAAGTACAGGACATCAACGAACTCCCTGTGTGAAAGG 301
DB 1651009 ATTAATTAAGTATGTTTAAAGTGTTAATTTAAGTGAACAGAACTATTAACAGAG 1650950
QY 302 AATTGATGCTGTTGAGTGAAGCGGCAATGCCCCCTAACCGAGCGGAAGTGTAGCTAG 361
DB 1650949 AATTTGTTTACGAAATTAAGACAGTATGTTATCTCGGATGAAATTTAAAGATTAG 1650890
QY 362 CCCAGATATCCGGGCGGCAATGTTGATATCTCGAAGACACCGTCAACATGAATGTT 421
DB 1650889 CAGATATTTTCCGCGCTCAAAATGATGATGAACCAAAAGTCTTACCAATTCATTA 1650830
QY 422 GGGGACCCCGGTTAAA 438

Db 1650829 GCGGTACAAATGATATA 1650813

RESULT 9
US-09-643-990A-1/c
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: P8186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 17.7%; Score 100.2; DB 4; Length 1830121;
Best Local Similarity 54.1%; Pred. No. 4,5e-21;
Matches 204; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 62 TCTCGTTTGTAGTGAAGTGAAGCGGAGTGAAGCGGATGCGGATGCGGATGTTGCC 121
DB 1651189 TATCGTTTCTACTGAAATGATGGAAGCATTTCCCGTGTGTTGTTATTTCTTC 1651130

QY 122 GCGGTGTTTAAACATGAGAGCTTGCGGTGCGGATGCGGAAACAGGGAGACGTTTCCC 181
DB 1651129 AACGCGCTTAATATGAAAGTTGACGGTTGACCAACGATGATCCACGCTTCTC 1651070

QY 182 GATACACATGCGTGGCGGGGATGAGAACACATGAGAACACTGACCAAGAACTCT 241
DB 1651069 GATATCACTTAAGCTGTGGCGATGCTCAGCGCTTGAGCAATCGAAAAACAATCTC 1651010

QY 242 ACAAGTGTAACTTAATTAAGTACAGACATCACGAAACTCCTGTGTGAGAAAGG 301

Db 1651009 ATAAATAGTGTATGTTTAAAGTGTAAATTTAAGTACGACGAAACATATGACGAG 1650950

QY 302 AATTGATGCTGTGAAGTGAAGCGCCCAATGCCCTTAACCGAGCGAAGTATGAGCTAG 361
DB 1650949 AATTTGTTTAGGAAAGTAAAGACAGTAGTTTCATCTGCGCATGAAATTAACGATTG 1650890

QY 362 CCCAGGTATTTCCGGGCGCCGATTTGTGATATCTCCGAAAGACACCGTCACATCGAATGCT 421
DB 1650889 CAGATATTTTCCCGGCTCAAAATTTGATGATGATACCAAAAGTCTTACAAATTCATTTAA 1650830

QY 422 GGGGACCCGGGTAAAA 438
DB 1650829 GCGGTACAAATGATATA 1650813

RESULT 10
US-09-540-236-676
Sequence 676, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA
FILE REFERENCE: 2709,2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 676
LENGTH: 546
TYPE: DNA
ORGANISM: M. catarrhalis
US-09-540-236-676

Query Match 16.0%; Score 90.2; DB 4; Length 546;
Best Local Similarity 53.9%; Pred. No. 1,2e-19;
Matches 256; Conservative 0; Mismatches 208; Indels 11; Gaps 3;

QY 52 AAACACACCCCTCTCTGTTTGTAGTGAAGTGAAGCGGAGTGTACTACCGCATTTGCCGA 111
DB 49 AAACATCTGATTTCTGTATTAATGAAAAGAAAGACAGGTTCCGTCTCTGTGTGCGGT 108

QY 112 CTATTTGCCCGCGGTGTTTAAACATTTGAGAGCTTGCGGTGGGTGCGGAAACAGGGG 171
DB 109 TTAATTTTCAGCGGTGCTAATATATTGATGACGCTCAATGTAGCACCGAGTGAACCA 168

QY 172 GACGTTCCCGATCAATGATGATGCGGCGGAGTGAAGAACCATGCAACATGCAACATGACC 231
DB 169 ACATCTCACGGCTGACTTTGACACACATTTACACACATGACAAATCGAACAAATCAC 228

QY 232 AAGCACTTCAAGTGTGTTAAGTAAATTAAGTACAGACATCCGA-----A 282
DB 229 AAGCAGCTGATTAACATGATGAGGTGTTAAAGTCCAAATCTGCCAGCTGTGACGC 288

QY 283 ACTCCCTGTGTGAAGGGAATGATGATGCTGTGCAAGTGAAGCGGCAATGCCCTAACCA 342
DB 289 GGTATCATGTTGAAGAGGTTGATGCTCATTTAAAGTGTGCAACGAGTTTAAACCGT 348

QY 343 GCGGAAGTATGAGTACGCCAGTATTTCCGGGCGCGCATTTGTGATATCTCCGAAGAC 402
DB 349 GAGGAATCAAGGTATATGCCGATATATTTCTGTCGCCAAATGTGATGTACGACAAAC 408

QY 403 ACCGTACCA-TGAATGTGTGGGAGACCGGGTAAATGTGACAAATCTCCAGATG-TT 460
DB 409 TTGTATACATTTTAAATTTACTGTGATGTGTGGAATTTGATGTTTATTCGATGTGATT 468

QY 461 GGCAGTGTGCAATTAAGAGGTGCTCGAAGCGGCAAAATGCTTTGTGTGCGGG 515
DB 469 GGTGTGACCGTATCTTAAGTGTATCGCTAGTGTATGTATGTCTGTG 523

RESULT 11
US-09-596-002-35
Sequence 35, Application US/09596002

Patent No. 6632636
GENERAL INFORMATION:
APPLICANT: Lagace, Robert, E.
APPLICANT: Paterson, Chandra
APPLICANT: Berg, Kim, L.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
FILE REFERENCE: RM-0008-4-US
CURRENT APPLICATION NUMBER: US/09/596,002
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: 60/140,121
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PERL Program
SEQ ID NO 35
LENGTH: 96109
TYPE: DNA
ORGANISM: M. catarrhalis
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incycle template ID No. 6632636 35
PUBLICATION INFORMATION:
US-09-596-002-35

Query Match 16.0%; Score 90.2; DB 4; Length 96109;
Best Local Similarity 53.9%; Pred. No. 1.8e-18;
Matches 256; Conservative 0; Mismatches 208; Indels 11; Gaps 3;

QY 52 AAACACACCCCTCTCTTTTGTGAAGATGAAACCCGAGTCTAACCCGCAATTGCCGA 111
DB 4168 AAACATCTGATTTCTGATTAATGAAAACGAAACAGATTCCTGCTGCTGCTGCTGCT 4227
QY 112 CTATTTGCCCGCCGCTGTTTAACTTGAAGAGCTTGGCGGTGGCGGGAACAGGGG 171
DB 4228 TTATTTCTGACCGGTGATTAATATTGATCGCTCAATGAGAGCAAGATGACCA 4287
QY 172 GACGTTTCCCGCATCACCATGCTGCTGCGCGGAGATGAAACACCATCGAACATGACC 231
DB 4288 ACGATCTCACGCGCTGATTTGACAAACATTAACAACATGACAAATCGAAATCAC 4347
QY 232 AAGCACTCTCAAGTGTGTTAACTTAATTAAGTACAGACATCAACCA-----A 282
DB 4348 AAGAGCTGCAATAACGATGAGTGTGTTAAAGTCCAAATCTGTCGAGCTGTACGGC 4407
QY 283 ACTGCTGTGTGAAAGGAAATGATGCTGTGAGAGTGAAGCGCAATGCCCTTAACCA 342
DB 4408 GGTATCATGTGTGAACAGAGTGTATGCTCATTTAAGTGTGTCACAGGGTCAAAACCGT 4467
QY 343 GCGAAGATGATGAGTGAAGCCAGTATTCGGGCGCCGCAATGATGATCTCCGAAGAC 402
DB 4468 GAAAGATTCAGCGTAAATGCGATATATTTGTCGCCAAATGATCGATGACACAAAC 4527
QY 403 ACCGTACCA-TCGAATGTGTGGGACCCGGGTAAATGTGTAGCAATCTCCAGATG-TT 460
DB 4528 TTGTATCAATTTAATTACTGTGTGATGTGCGCAATTTGATGATGATGATGAT 4587
QY 461 GCGCAATGTGCATTAAGAAGTGTGTCGAACGGGCAAAATGCTTTGGTGGCGGG 515
DB 4588 GGTGTGACCGTATCTTAAGAGTGTGACGCTCAAGTGTGATGTGATGCTGTG 4642

RESULT 12
US-09-988-1/c
Sequence 1, Application US/09790988
Patent No. 6632935
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: MATANABE, HIDEKI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENE OF BACTERIAL SYMPTOM OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 13.8%; Score 78; DB 4; Length 640681;
Best Local Similarity 50.0%; Pred. No. 4.8e-14;
Matches 195; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 49 ATGAAACACACCCCTCTCTTTTGTGAAGATGAAACCCGAGTCTAACCCGCAATTGCC 108
DB 248814 ATGCAAGATATTTATCTATTTTATGGAATGATGATGATGATGATGATGATGATGAT 248755
QY 109 GGACTATTTGCCCCCGCTGTTTAACTTGAAGAGCTTGGCGGTGGCGGGAACAG 168
DB 248754 GGACTCTTTTCAAGAGGATATATATGAACTATTTAGCTAGCACTACTGAAAT 248695
QY 169 GGGAGCGTTTCCCGCATCAAGTGTGTCGCGGAGTGAACACCATGAAACAATG 228
DB 248694 CCTACTTATCTAAATGATTAACAAACAGTAAATGAAAGATTAATGAAATTA 248635
QY 229 ACCAAGCAACTCTAAGTGTGTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 288
DB 248634 GAAAGCAAGTCAATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 248575
QY 289 TGTGTGAAAGGAAATGATGCTGTGAGAGTGAAGCGCAATGCCCTTAACCGAAG 348
DB 248574 CATATGAAAGCGGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 248515
QY 349 GTGATGAGTACAGGATTTCCGGGCCGCTTGTGATATCTCCGAAGACACCGTC 408
DB 248514 ATAAACATATTAATGAAATTTGAGAGGCAAAATTTGACATTAATGATTAAT 248455
QY 409 ACCATGATGATGATGAGGAGCCCGGTAA 438
DB 248454 ATATTGCAATTGTCTGTACTACGAAAAA 248425

RESULT 13
US-08-961-527-8
Sequence 8, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373

```
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6211 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-8

Query Match      13.2%; Score 74.8; DB 4; Length 6211;
Best Local Similarity 50.9%; Pred. No. 4,8e-14;
Matches 230; Conservative 0; Mismatches 217; Indels 5; Gaps 2;

QY 72 AGTGAAGATGAGCGGAGTGTCTAACCGGATTCGGGACTATTTGCCCGCGTGGTTT 131
DB 1802 ACTACAAATGCTTAGAGAGTCTCATGCTTTACAGGTGTCCTATCTGTCGTACAGT 1861
QY 132 TAACTTGAAGAGCTTGGCGGTGGGTCCGCGAACAAGGCGGACGTTTCCCGCATCACCAT 191
DB 1862 TAACTTGAAGAGCATCTGTTGGAGCAACAAGATCCGAATGATCGGTATCACTAT 1921
QY 192 GGTGTGCCGGGGAGT--AGAACACATGGAACAACTGACCAAGCACTCTCAAGTT 248
DB 1922 TATTATGATGTGCTTCTCATGATGAAGTGGAGCAATCAACAACGCTCAATCGTCA 1981
QY 249 GATTACCGTAAATTAAGTACAGACATCACCGAACTCCCTGTGTGAAAGGAGATTGAT 308
DB 1982 GATTGATGTGATTCGATTCAGATATACAGACAGCTCATTTGGAGCCGAGGTGAT 2041
QY 309 GCTGTGAAGTGAAGCCCAATGCCCTTAACCGAGCGGAAGTGAATGAGCTAGCCCAAGT 368
DB 2042 TTTGTGTAAAGTGTGAGGCGCAGCTGAGAAAGAGCTGAGATTTTAAAGGATTAATCAAC 2101
QY 369 ATTCGGGCGCCGATTTGATATCTCGAAGACACCGTCAACATCGAATGTTGGGGGAC 428
DB 2102 TTTCGTCGAACAGTAGTAGAGCTAGCCCAAGCTCGATTTCCATTCAGATGACGGGAAA 2161
QY 429 CCGGGTAAATGTGA--GCAATCTCTCAGATGTTGGCCCAAGTGGCATTTAAAGAGTGC 486
DB 2162 TGCAAGAAAGAGCGAAGCCCTATTGCGAGTCATTCGCGCATTCGCAATATGTC 2221
QY 487 TCGAAGCGGCAAAATGCTTTGTCGTGCGGGAAAT 518
DB 2222 TCGAAGCGGTGCACTGAGATTATCCCGCGAAT 2253
```

```
RESULT 14
US-08-916-421B-1
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Buit et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB375
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28252)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
```

```
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (103988)..(103988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
```

```

LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (773455)..(773455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)

```

```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

```

```

Query Match      12.4%; Score 70; DB 4; Length 1664976;
Best Local Similarity 51.1%; Pred.No. 3.3e-11;
Matches 240; Conservative 0; Mismatches 225; Indels 5; Gaps 3;

```

```

QY 55 CACACCCCTCTGTTTGTAGTGAAGATGAACCCGAGCTTAACCCGATTCGGGACTA 114
DB 1307221 CATTGATCTCCGCTGATGTTTAATAAACCAGAGATTGACAGAGATTCCAGGTTA 1307270
QY 115 TTTCGCCGCGTGTGTTTAACTAGAGAGCTTGCGGTGCGCGGAACAG393GAC 174
DB 1307271 TTACAGAGAGAGGGTTTAATTAATTCAMGTAATACGTCGGAATTAACAGAAATCCACAA 1307330
QY 175 GTTCCCGCATCAACATGCTGTCGCGGGGATGAGAACACCATCGAACTGACCAAG 234
DB 1307331 ATTTCAAGAGTTACGATGTTGTTAATGAGATGATTAAGATTTAGAGCAGTTATCAAA 1307390
QY 235 CACTCTCAAGTTGTTAGCTAATTTAAAGTACAGACATCACCGAAATCCCTGTGTG 294
DB 1307391 CACTCAACAATTAATTAAGTTAATTAAGTTAGTTAGAGAAAATAATCCGTT 1307450
QY 295 GAAAGGAATTGATCTGCTGTAAGTGACG---CCAATGCCCTTAACCGAGCGAAATG 351
DB 1307451 CAGAGAGAGCTCTGTTTAAATTAAGATTTATCACCAACAGAGAGTGAATAATCAAGTT 1307510
QY 352 ATGAGTAGCCAGCTTATTCGCGGCGCCGCAATGTGATCTCCGAAGACACCGTCACC 411
DB 1307511 ATTCAATTAACAAGATTTTAGAGAAATTTGATGATTTAATCCAGAAATCTTTAATT 1307570
QY 412 ATCGA-ATGCTGGGGAGCCCGGTTAAATGTTAGCAATCTCCAGAT-GTTGGCAAATT 469
DB 1307571 GTAGAGATACTGTAAGTAAGATAAATAACCATTAATTAAGTTAATCACTTA 1307630
QY 470 GGCAATTAAGAGTGGCTCGAAGCGGCAAAATTCCTTGTGGCGGGAATC 519
DB 1307631 GGAATTAAGAAATGCAAGACTGGAATTAATCTCCTTAAGCAAGGGAATC 1307680

```

```

RESULT 15
US-09-426-568A-3
Sequence 3, Application US/09426568A
Patent No. 6348643
GENERAL INFORMATION:
APPLICANT: Kakeluda, Genichi
APPLICANT: Costello, Colleen
APPLICANT: Sun, Ming
APPLICANT: Hu, Weiming
TITLE OF INVENTION: Genes and Vectors for Confering Herbicide Resistance
FILE REFERENCE: 008103/195497
CURRENT FILING DATE: US/09/426,568A
PRIOR APPLICATION NUMBER: 1999-10-22
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 4895
TYPE: DNA
ORGANISM: Arabidopsis sp.
FEATURE:
NAME/KEY: Promoter
LOCATION: (1)..(757)
OTHER INFORMATION: Promoter Region
NAME/KEY: misc_feature

```

```

; LOCATION: (717)
; OTHER INFORMATION: Transcriptional Starting Point
; NAME/KEY: misc_signal
; LOCATION: (758)..(760)
; OTHER INFORMATION: Start Codon
; NAME/KEY: misc_signal
; LOCATION: (4737)..(4739)
; OTHER INFORMATION: Stop Codon
; OTHER INFORMATION: n at position 694 can be a, c, g, or t
US-09-426-568A-3

```

```

Query Match      10.9%; Score 61.6; DB 4; Length 4895;
Best Local Similarity 58.7%; Pred. No. 8.8e-10;
Matches 128; Conservative 0; Mismatches 84; Indels 6; Gaps 1;

```

```

QY 49 ATGAACAACACCCCTCTCTGTTTAAAGTGAAGTGAAGCCGGAGTGTAAACCCGATTGCC 108
DB 1088 AGAAGACACACATTTTCAATTTGTGTGAGCGAAGCGAATGATTATAGATTGCA 1147
QY 109 GGAATATTTGCCCCGCGTGTGTTTAACTGAGAGCTTGGCGGTGGGTGCGCGGAACAG 168
DB 1148 GGAAGTTTGCAGAGGAGGATACATATATGAGAGTCTTGTGTGCTG-----AAC 1201
QY 169 GGGGACGTTTCCCGCATCACCATGCGTGGCGGGGATGAGAACACCATGAAACAACTG 228
DB 1202 AGAGCAAGAGGCTCTATTCACCATAGTTGTGTGSACTGAAGGTACTTCAGCAGGTC 1261
QY 229 ACCAAGCAACTCTACAAGTTGTTAAGTTAAAGT 266
DB 1262 ATCGAGCAACTCGAGAGCTCGTTATGTCTAAAGT 1299

```

Search completed: July 25, 2004, 12:54:57
 Job time : 54.0987 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 07:28:18 / Search time 272.203 Seconds

(without alignments)
10146.437 Million cell updates/sec

Title: US-09-893-033-17

Perfect score: 565

Sequence: 1 gtggaatttaccacatg.....ccctggaatccagttttag 565

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3216467 segs, 244419694 residues 6432934

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US08A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141.2	25.0	1721	9	US-09-732-618-17
2	137	24.2	504	13	US-10-282-122A-26262
3	137	24.2	507	13	US-10-282-122A-28638
4	135.8	24.0	504	13	US-10-282-122A-25559
5	135.4	24.0	1813	9	US-09-732-618-9
6	134.6	23.8	489	13	US-10-282-122A-13980
7	133.8	23.7	492	13	US-10-282-122A-29826
8	132.8	23.5	1297	9	US-09-732-618-5
9	132.2	23.4	489	13	US-10-282-122A-29150
10	129.6	22.9	525	15	US-10-156-761-7719
11	129.6	22.9	9025608	15	US-09-732-618-15
12	129	22.8	1435	9	US-09-732-618-15
13	127.6	22.5	489	13	US-10-282-122A-15104
14	126.6	22.4	489	13	US-10-282-122A-13405

15	126	22.3	525	13	US-10-282-122A-11953	Sequence 11953, A
16	132.8	21.7	1525	13	US-10-424-599-23293	Sequence 23293, A
17	117.2	20.7	510	13	US-10-282-122A-27732	Sequence 27732, A
18	116.6	20.6	1561	17	US-10-437-963-3796	Sequence 3796, Ap
19	114	20.2	2256646	17	US-10-470-565-1	Sequence 1, Appl
20	112	19.8	960	9	US-09-770-445-311	Sequence 311, Ap
21	112	19.8	1673	9	US-09-997-900-1	Sequence 7946, Ap
22	111.8	19.8	492	9	US-09-815-242-7946	Sequence 30523, A
23	111.8	19.8	1861	9	US-10-282-122A-30523	Sequence 23, Appl
24	110	19.5	1861	9	US-09-732-618-23	Sequence 3458, Ap
25	109.4	19.4	516	9	US-09-738-626-3458	Sequence 1, Appl
26	109.4	19.4	3309400	9	US-09-738-626-1	Sequence 5844, Ap
27	108.4	19.2	1464	17	US-10-437-963-5844	Sequence 7233, A
28	105.8	18.7	519	13	US-10-282-122A-7233	Sequence 2330, A
29	105.8	18.7	519	13	US-10-282-122A-3330	Sequence 2950, Ap
30	105.6	18.7	556	17	US-10-021-323-2950	Sequence 17529, A
31	104.4	18.5	522	13	US-10-282-122A-17529	Sequence 1987, Ap
32	104.2	18.4	489	9	US-09-815-242-3987	Sequence 7645, Ap
33	104.2	18.4	522	9	US-09-815-242-7645	Sequence 31295, A
34	103.8	18.4	489	13	US-10-282-122A-31295	Sequence 9033, A
35	103	18.2	489	13	US-10-282-122A-9034	Sequence 41304, A
36	101.8	18.0	492	13	US-10-282-122A-30933	Sequence 7159, Ap
37	100.6	17.8	455	13	US-10-282-122A-41304	Sequence 1, Appl
38	100.2	17.7	452	9	US-09-815-242-7159	Sequence 22364, A
39	100.2	17.7	492	13	US-10-282-122A-22364	Sequence 12917, A
40	100.2	17.7	1830121	15	US-10-329-960-1	Sequence 1, Appl
41	100.2	17.7	1830121	16	US-10-329-960-1	Sequence 12917, A
42	99.6	17.6	2130	13	US-10-424-599-129217	Sequence 41728, A
43	97.6	17.3	495	13	US-10-282-122A-41728	Sequence 33628, A
44	97.4	17.2	489	13	US-10-282-122A-33628	Sequence 38994, A
45	95.8	17.0	492	13	US-10-282-122A-38994	

ALIGNMENTS

RESULT 1
US-09-732-618-17
Sequence 17, Application US/09732618
Patent No. US2001004939A1
GENERAL INFORMATION:
APPLICANT: Abbell, Lynn
APPLICANT: Falco, Carl
APPLICANT: Farnoch, Omolayo O.
TITLES OF INVENTION: Small Subunit of Plant Acetolactate Synthase
FILE REFERENCE: BB1435 US NA
CURRENT APPLICATION NUMBER: US/09/732,618
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/174,437
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Microsoft Office 97
SEQ ID NO 17
LENGTH: 1721
TYPE: DNA
ORGANISM: Glycine max
US-09-732-618-17

Query Match 25.0%; Score 141.2; DB 9; Length 1721;
Best Local Similarity 58.2%; Pred. No. 4.2e-37;
Matches 285; Conservative 0; Mismatches 203; Indels 2; Gaps 2;

QY	54	ACACACCCCTCTGTTTGTAGTGAAGAGCGGAGGCTACCCGACT	113
DB	1005	ACACACCTTATCATGCTTGTGACATGCTCTGAGATTTAAATTTGACAGAG	1064
QY	114	ATTGCCCCCGGCTGTTTAACTAGAGCTTGGCGGTGCGCGGACAGGGGA	173
DB	1065	TTTGTCTAGAGAGGCGTATACATTCAGAGTTAGCTGATGACATGACGAGG	1124
QY	174	CGTTCCCGATACCATGCTGTCGGCGGATGAGACACATGACCACTGACCA	233
DB	1125	ACTTTCGACTTACACTGTGTCTCTGGGACAGATGATGATTAAGCAAGTTGTGA	1184

ORGANISM: Mycobacterium bovis
US-10-282-122A-26262
Query Match 24.2% Score 137; DB 13; Length 504;
Best Local Similarity 57.7% Pred. No. 6.3e-36;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;
Db 1185 GCAACTGTATAGCTAGTAGAGCTACATGAGGTTCTGATATACACCACTTGCATTGGC 1244
Qy 234 GCAAAGGAAATTGATGCTGCGTGAAGGTGAGGCCCAATGCCCTTAACCGAAGATGAT 353
Db 1245 TGACCGAAGATTGATGCTAAATTAAGATTGCTTAAAGTGTGACAGCGGTGATGCTT 1304
Qy 354 TGACCTAGCCCAAGTATCCGGGCCCGCATTTGATGATATCTCCGAAACACCGTACCAT 413
Db 1305 TGAATTTGCTAGCATTTTTCGGGGCTAAAGCTGTGATGATCTGATCACAATTAAGTCT 1364
Qy 414 CGA-ATGCTGGGGGACCCGGGTAAATGTAGCATCTCTCAATGTT-GGCCAAGTTGG 471
Db 1365 GGAGCTTACTGAGATTTTGAACAAGATGGTTGCAATTCAGAGATTGTAACCATATGG 1424
Qy 472 CATTAAAGAGGTGCTCGAACCGGCAAAATGCTTTGTGCGGGAAATCCGGCGTCAATAC 531
Db 1425 CATTTGAGGTGCGACGAACTGGGGCAATGCTCTAGTCCGCGAGTCCGGTGTGACATC 1484
Qy 532 GGAATATCTG 541
Db 1485 CAAGTACTTG 1494
RESULT 2
US-10-282-122A-26262
; Sequence 26262, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 26262
; LENGTH: 504
; TYPE: DNA

ORGANISM: Mycobacterium bovis
US-10-282-122A-26262
Query Match 24.2% Score 137; DB 13; Length 504;
Best Local Similarity 57.7% Pred. No. 6.3e-36;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;
Db 41 TCCCCCATGAAACACACCTCTCTGTGTTTAAAGTGAAGAGACCGGAGTGTAAACC 100
Qy 2 TGAAGCCGAAAGCGACAGCTTGTGAGTGTGCGAAGACAGCCGCGTGTGCGGC 61
Db 101 GCATTGCGGACATATTTGCTCCGCGCGTGTGTTTAAATTGAGAGCTGCGGTGGGTGCG 160
Qy 62 GGGTGGCGCGCTGTCTCCGGGCGGGTTTCAACATGAGTGTGGCGGTGGGTGCA 121
Db 161 CGAAACAGGGGACGTTTCCCGCATCACATGCTGTGTCCGGGGATGAAACACATCG 220
Qy 122 CCGAGTGAAGACAGGTCAGGATGACGATCGGTGTTCCGCGAGAGCATCCGCTCG 181
Db 221 AACCACTGACCAAGCACTCTACAAAGTTGTTAAAGTAAATTAAGTACAGGACATCACCG 280
Qy 182 ACAGATCACCAAGCACTCTACAAAGTTGTTAAAGTAAATTAAGTACAGGACATCACCG 241
Db 281 AAATCTCTGTGTGGAAGAGGAAATGATGCTGTGAGGTAGAGGCCCAATGCCCTTAAC 340
Qy 242 ACAGACATCGGTGTGTCAGGGAAATGGCGCTCATCAAGTCAAGCCGACCGCGAGCC 301
Db 341 GAGCGGAAGTGAATTAGGCTAGCCAGGATTCGGGGCCCGCATTTGATATCTCCGAAG 400
Qy 302 GAGCGCAAGTGAATGCAAGCGGTGAATCTGTTTCCGCCCAAGTATGACTATCCCGG 361
Db 401 ACACCGTACCACTGCAATGATGTGGGGACCG-GGTAAATGTAGCAATCTCCAGATGT 459
Qy 362 AGTCATTGACCGTTCAGGCCACCGGTAAACCGCGCAAGTGAAGGCCCTGCTGGGGTGT 421
Db 460 TGGCCAAATT-GGCAATTAAGAGTGTGCTCAACCGGCAAAATTTGTTGTGGGGGAT 518
Qy 422 TGAAGCCGTTGCGTATTCGCAAAATCGCCAAATCGGAATGTGTGCTGTCCGGGTG 481
Db 519 CCGGCTCA 527
Qy 482 CGCGCGCA 490
RESULT 3
US-10-282-122A-26262
; Sequence 26262, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 28638
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28638

```

```

Query Match      24.2%; Score 137; DB 13; Length 507;
Best Local Similarity 57.7%; Pred. No. 6,3e-36;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;

```

```

QY 41 TGGCCCCCATGAACACACCCCTCTGTTTGAAGATGAACCCGAGTCTAACCC 100
DB 2 TGAAGCCCGAAGACGACACGTTGTGGTGTGTAAGACAGCCCGCTGCTGCGC 61
QY 101 GCATTGCCGACTATTTCGCCCGCGTGTGTTTAACTTGAAGCTTGCCTGCGGTGCG 160
DB 62 GGGTGGCGGCGCTGTTCTCCGCGCGGTTTCAATGAGTGGTGGCGGTGGTCCA 121
QY 161 CGGAACAGGGGAGCGTTTCCCGCATCACCATGTGTGTCGGGGGATGAAACACCATCG 220
DB 122 CCGAGTGCAGAGCAGGTGACGATGACCATGCTGTCTCCGCCAGACACTCCGCTCG 181
QY 221 AACCACTGACCAAGCAACTCTTCAAGTTGTTTAACTTAAATTAAGTACAGACATCACCG 280
DB 182 AGCAGATCACCAAGACAGCTCAACAAGCTGATCAAGTATCTCGAGACAGACG 241
QY 281 AAATCTCCTGTGTGAAAGGGAATTGATGTCTGTGAAGGTGAGCGCCCAATGCCCTTAA 340
DB 242 AGGACACATCTCGTGTGACCGGAATTGGCGCTCATCAAGCTCCAGCCGCGCAGCC 301
QY 341 GAGCGGAAGTGAATGAGTGAAGCCAGGTATTCCGGGCGCGCATTTGTGATATCTCCAG 400
DB 302 GGAAGCAATGATTCGAAGCGGTGAATCTGTTTCGCGCAAGTATTAAGTATCTCCGCG 361
QY 401 ACACCGTCAACATGCAATGTGTGGGAGCCCG-GGTAAATGATGAACAATCTCCAGATGT 459
DB 362 AGTCAATTGACCGTCCAGGCGCACCGGTAAACGGGCGCAAGTTAGAGGCCCTGCTGCGGGGT 421
QY 460 TGGCCAAAGT-GGCATTAAAGAGGTGCTCGAACGGGCAAAATTCCTTGTGTGGGGAAT 518
DB 422 TGAAGCCGTTGCGTATTCGCGAATTCGCCAATCCGGAATGATGTGCTGTCCGCGGTTC 481
QY 519 CGGCGGTCA 527
DB 482 CGCGCGGCA 490

```

```

RESULT 4
US-10-282-122A-25559
; Sequence 25559, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith

```

```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 25559
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25559

```

```

Query Match      24.0%; Score 135.8; DB 13; Length 504;
Best Local Similarity 56.0%; Pred. No. 1.6e-35;
Matches 277; Conservative 0; Mismatches 217; Indels 1; Gaps 1;

```

```

QY 41 TGGCCCCCATGAACACACCCCTCTGTTTGAAGATGAACCCGAGTCTAACCC 100
DB 2 TGAAGCCCGAAGACGACACGCTGTGGTGTGTAAGACAGCCCGGTGCTGCGC 61
QY 101 GCATTGCCGACTATTTCGCCCGCGTGTGTTTAACTTGAAGCTTGCCTGCGGTGCG 160
DB 62 GGGTGGCGGCGCTGTTCTCCGCGCGGTTTCAATGAGTGGTGGTGGTCCA 121
QY 161 CGGAACAGGGGAGCGTTTCCCGCATCACCATGTGTGTCGGGGGATGAAACACCATCG 220
DB 122 CCGAGTGCAGAGCAGGTGACGATGACCATGCTGTCTCCGCCAGACACTCCGCTCG 181
QY 221 AACCACTGACCAAGCAACTCTTCAAGTTGTTTAACTTAAATTAAGTACAGACATCACCG 280
DB 182 AGCAGATCACCAAGACAGCTCAACAAGCTGATCAAGTATCTCGAGACAGACG 241
QY 281 AAATCTCCTGTGTGAAAGGGAATTGATGTCTGTGAAGGTGAGCGCCCAATGCCCTTAA 340
DB 242 AGGCAACTCTGCTGTGCGGGAATTGGCGTATTAAGGTGCGGCTCGACGCCGACCC 301
QY 341 GAGCGGAAGTGAATGAGTGAAGCCAGGTATTCCGGGCGCGCATTTGTGATATCTCCAG 400
DB 302 GAGTCAAGTCAATGCAAGGGGTGAACCTGTTTCGCGCAAGTATTAAGTATCTCCGCG 361
QY 401 ACACCGTCAACATGCAATGTGTGGGAGCCCGGTTAAATGATGAACAATCTCCAGATGT 459
DB 362 AGGCGCTGAGTGAAGAGCGCACCGGTGACCGGCAAGATCGAGGCGCTGTGCGGGTGC 421
QY 460 TGGCCAAAGTGTGATTAAGAGGTGCTCGAACGGGCAAAATTCCTTGTGTGGGGAATC 519

```

```

; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPIITA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/151,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,536
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13980
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-13980
Query Match      23.8%; Score 134.6; DB 13; Length 489;
Best Local Similarity 59.1%; Pred. No. 4.1e-35;
Matches 230; Conservative 0; Mismatches 159; Indels 0; Gaps 0;
49 ATGAACACACCCCTCTCTGTTTATGTTGAAGATGAGCCGAGTGTCTAACCCCGATTGCC 108
Db 1 ATGAGACACATCATTTCCGTCTGCTGAAATGAACCGGACGCGTGTGCGCGTGTGTC 60
QY 109 GGACTATTTCGCCGCGCGTGTATTAACATTGAGAGCTTTGCGGCGTGGCGGAGAC 168
Db 61 GGTCTGTTCTCGGACCGCGCTACAAATCGAAACCTTGAACGTTGGCGCGGACCGAAGAC 120
QY 169 GGGAGCGTTTCCGCATACACCATGCTGTGCGCGGAGTGAACACACCATGGAACATG 228
Db 121 CAGTCGCTGTCGCGCTGACCATGCTGTCCATGCGGCTCCGACACAGTATCGAACAATC 180
QY 229 ACCAAGCACTCTTCAAGTTGTGTTAACTGATATTTAAATAAGAGACATCACCGAACTCC 288
Db 181 ACCAAGCATCTGAACCGCTGATCGAGTGTGATGAATGTGTGAGCTCTGACCGACGGTCA 240
QY 289 TGTGTGAAAAGGGAATGATGCTGTGTAAGGTGAGCCCATGCGCCCTTAACCGAGCGGAA 348
Db 241 CACATCGACGGCGAGCTGATGCTCATCATGAAGTGAAGGCAATGGGCAAGGAGCGGAAGAA 300
QY 349 GTGATTAGCTAGCCCAAGTATTTCGGGCGCCGATTGTGATATCTCCGAAGACCGCTC 408

```

Db 301 ATGAAGCGGATGGCGGACATTTCCGGGCGGCATCATCATGACGAGAAAGACCTAC 360
QY 409 ACCATCGAATGTGGGGGACCCGGGTAA 437
Db 361 ACGATCGAATTGACGGGGCGAGCGACAA 389

RESULT 7

US-10-282-122A-29826
; Sequence 29826, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyeth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR FILING DATE: 2003-02-20

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

PRIOR FILING DATE: 2000-03-21, 078

QY 229 ACCAACAACCTCTACATAGTTGTTAAGTAATTAAGTACAGACATCACCGAACTCC 288
Db 181 ACCAACAACCTCTACATAGTTGTTAAGTAATTAAGTACAGACATCACCGAACTCC 240
QY 289 TGTGTGGAAGGGAATTTGCTGTGTAAGGTGAGCGCCCAATGCCCTTAACCGAGCGAA 348
Db 241 TTGTGTGGAAGGGAATTTGCTGTGTAAGGTGAGCGCGCCCAATGCCCTTAACCGAGCGAA 300
QY 349 GTGATGAGCTAGCCAGGATTTCCGGGCGCCGATTTGATATCTCCGAGACACCGTC 408
Db 301 TTTTACGCTTACCGAATCTACCGGGGACAGCATCATGACGTAAACGACCGAGCTAT 360
QY 409 ACCATCGAA 417
Db 361 ACGATTGAA 369

RESULT 8

US-09-732-618-5

; Sequence 5, Application US/09732618

; Patent No. US20010044939A1

; GENERAL INFORMATION:

APPLICANT: Abell, Lynn

APPLICANT: Ramodu, Omolayo O.

TITLE OF INVENTION: Small Subunit of Plant Acetolactate Synthase

FILE REFERENCE: B31435 US NA

CURRENT FILING DATE: 2000-12-08

PRIOR FILING DATE: 2000-12-08

PRIOR FILING DATE: 2000-01-04

NUMBER OF SEQ ID NOS: 43

SOFTWARE: Microsoft Office 97

SEQ ID NO 5

LENGTH: 1297

TYPE: DNA

ORGANISM: Zea mays

US-09-732-618-5

Query Match 23.5%; Score 132.8; DB 9; Length 1297;

Best Local Similarity 57.2%; Pred. No. 2.7e-34;

Matches 279; Conservative 0; Mismatches 207; Indels 2; Gaps 2;

QY 55 CACACCTCTCTGTTTAACTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 114

Db 542 CATAACCTCTCTCAATCCTTGAATGATGTGCTGAGTGTCTCAACTTGTAAACAGAGCTC 601

QY 115 TTGCGCGCGCTGTTTAACTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 174

Db 602 TTGCTGCGAGGCGGTAACTATATAGAGAGGCTTGTGCTGAGCTGAGAGGAGG 661

QY 175 GTTTCGCGATCACTGATGCTGCGGAGGAGTGAAGAACCATTCGAACACTGACCAAG 234

Db 662 ATTTCAGTATTCACACAGTGTCTCTGATCTGTGAATTCATTGGAGTTAGTTCAG 721

QY 235 CAACCTACAGTGTGTTAACTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 294

Db 722 CAGCTTACAGTGTGTTAACTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 781

QY 295 GAAAGGGAATGATGCTGTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 354

Db 782 GAAAGGGAATGATGCTGTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 841

QY 355 GAGCTAGCCAGGATTTCCGGGCGCGATTTGATATCTCCGAGACACCGTCA-CCAT 413

Db 842 GATATGCTGAATTTCCGAGCAAAACCATTTGATATCTCCGAGACACCGTCA-CCAT 901

QY 414 CGAATGCTGGGGAACCGGGAATTTGATATCTCCGAGACACCGTCA-CCAT 472

Db 902 CAGCTTACAGTGTGTTAACTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 961

QY 473 ATTAAAGGAGTGTGTTAACTTGAAGTGAAGCGGAGTGTAAACCGCATTTGCCGAGCTA 532

Db 962 ATCTGGAGGTCGCCAGACTGACGAGTGGCACTGGTCGTGAATCGAAGTGCATCC 1021
QY 533 GAATATCT 540
Db 1022 AAGTACT 1029

RESULT 9

US-10-282-122A-29150
; Sequence 29150, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 29150
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-29150

Query Match 23.4%; Score 132.2; DB 13; Length 489;
Best Local Similarity 59.9%; Pred. No. 2.7e-34;
Matches 221; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
QY 49 ATGAACACACCCCTCTGTTTGAAGATGAACCGGAGTCTAACCCGCAATGGC 108
Db 1 ATGGACATATCTTATCTCTTCTGATGAAACGAATCAGGTGAGTACCCCGTGTG 60
QY 109 GAGCTATTTGGCCCGCGGTGTTTAACTTGAAGAGCTTGGCGGTGGGTGGGGAACG 168
Db 61 GGTGTGCTCTGACCGCGATTAACATATGATTCCTTGGCGGTGGCGCGAAGAC 120
QY 169 GGGGACGTTTCCCGCATCACGATGAGTGGCGGGGGATGAGAACACGATCGAACATG 228
Db 121 AAAACCTGTGACGAGTACCATGTTTACCAAGCGGAGGATGATCGAACAAATC 180

QY 229 ACCAAGCACTCTACAGTTGGTTAAGTAAATTAAGACAGACATCACGAAACTGCC 288
Db 181 ACCAAGCACTCAATAATTAATTAAGTAAAGTGTGATTTTAAAGAAAGCCGT 240
QY 289 TGTGTGAAAGGAATTAATGCTGTGTAAGTGAAGGAGCGCAATGCCCTTACCGAGCGGA 348
Db 241 TTGTGTAAGCGGAATGATGTTGTATAAAGTCCGTGCGCGGCAAGAACCGGACGAA 300
QY 349 GTGATTGAGCTAGCCGAGTATTCGGGCGCGCATTTGATGATCTCGGAAGACCGCTC 408
Db 301 TTTTACGCTTACCGAAATCTACCGGGGAGCATCATCGATTAACCGACCGCAGCTAT 360
QY 409 ACCATCGAA 417
Db 361 ACCATTGAA 369

RESULT 10

US-10-156-761-2719
; Sequence 2719, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIRA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO: 2719
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(525)
US-10-156-761-2719

Query Match 22.9%; Score 129.6; DB 15; Length 525;
Best Local Similarity 58.2%; Pred. No. 2.2e-33;
Matches 284; Conservative 0; Mismatches 199; Indels 5; Gaps 3;
QY 52 AAACACACCTCTCTGTTTGAAGATGAACCGGAGTCTTAACCCGCAATGGC 111
Db 7 AAGCACACCTCTCTGCTCTGTCGAGAAACAAGCCGGTGTCTCGCCGGATCACCGCC 66
QY 112 CTATTTGGCCCGCGGTGTTTAACTTGAAGAGCTTGGCGGTGGGTGGGCGGAACAGGG 171
Db 67 CTGTTCTCCGCGCGGGGTTTAACTTGAAGAGTCTGCTCGGTGTGTACCGAGACACCC 126
QY 172 GACGTTTCCCGCATCACCATGATGATGATGATGATGATGATGATGATGATGATGATG 228
Db 127 GACATCTCCCGCATCACCATGATGATGATGATGATGATGATGATGATGATGATGATG 186
QY 229 ACCAAGCACTCTTACAGTTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 288
Db 187 ACCAAGCACTTACAGTTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 246
QY 289 TGTGTGAAAGGAATTAATGCTGTGTAAGTGAAGGAGCGCAATGCCCTTACCGAGCGGA 348
Db 247 GCGGTGACGCGCAATCTGTTCTGTGTAAGTGAAGGAGCGCGCGCAACGAGCGCTCCAG 306
QY 349 GTGATTGAGCTAGCCGAGTATTCGGGCGCGCATTTGATGATCTCGGAAGACACGCTC 408

Page 7

QY	409	ACCATGGA-ATGATGGGGGACCCGGGTAATAATGATGCAATCCCTCCAGATGTTGGCAAG	467
Db	3354047	ACCATGAGAGCCACCGGCGGACGACACAACTGTGAGGCCATGCTCAGATGATGTAAGCCG	33539888
QY	468	TT-GGCAATTAAAGAGGTGCTCGAACGGGCAAAATTGCTTTGTGTGGGGAATCCGGCGTTC	526
Db	3353987	TTGGGATCAAGAGAGCTGCTCCAGTCGGGAACGATCGGATCGGACGTGGGGCACGTTTCG	33539288
QY	527	AATACGGA	534
Db	3353927	ATCACGGA	3353920

	RESULT 12	
	US-09-732-618-15	
	; Sequence 15; Application US/09732618	
	; Patent No. US20010044539A1	
	; GENERAL INFORMATION:	
	; APPLICANT: Abell, Lynn	
	; APPLICANT: Falco, Carl	
	; APPLICANT: Famodu, Omolayo O.	
	; TITLE OF INVENTION: Small Subunit of Plant Acetolactate Synthase	
	; FILE REFERENCE: B1435 US NA	
	; CURRENT APPLICATION NUMBER: US/09/732,618	
	; CURRENT FILING DATE: 2000-12-08	
	; PRIOR APPLICATION NUMBER: 60/174,437	
	; PRIORITY FILING DATE: 2000-01-04	
	; NUMBER OF SEQ. ID NOS: 43	
	; SOFTWARE: Microsoft Office 97	
	; SEQ ID NO 15	
	; LENGTH: 1435	
	; TYPE: DNA	
	; ORGANISM: Oryza sativa	
	US-09-732-618-15	
QY	Query Match	22.8%; Score 129; DB 9; Length 1435;
	Best Local Similarity	56.6%; Pred. No. 5.7e-33;
	Matches 277; Conservative	0; Mismatches 210; Indels 2; Gaps 2;
DG	54 ACAACCCCTCGTGGTTTAAGTAGGAAGCGAGGTGCTAACCCGCATTCGCCGACT	113
Dg	629 ACATACCTCATTCACCTCCTTGCAATGATTGCCCTGGGTTCTCAACAATTGTACAGGGGT	688
YY	114 ATTGCCCCGCGTGGTTTTAACTAGAGCCTGGCGGTGGGAGTGGCGGAACAAGGAGGA	173

Query Match	22.8%	Score 129	DB 93	Length 1435
Best Local Similarity	56.6%	Pred. No. 5.7e-33		
Matches 277	Conservative 0	Mismatches 210	Indels 2	Gaps 2
QY	ACAACACCTCTCTGTTTACTTGAAGAATGAAGCCGAGAGTCTAACCCGCAATTCGCGACT	113		
Db	ACATACCTCTATCCATCCTTGTCAATGATGTCCTCGGGTGTCTCAACATTGTTACAGGGGT	688		
QY	114 ATTTGCCCCCGTGGTTTAAACATTGAGACTTGGCGGTGGGTGGCGGCAACAGGGGGA	173		
Db	689 CTTTGCTCGAGAGGCTACATATACAGAGTCTTGCTGTAAGCCCACTGAAAAGTCAAG	748		
QY	174 CGTTTCCCGCATCACCATGTGTGTCGGGGGAGTAGAACAACCATTCACAACATGACCAA	233		
Db	749 CTTTGGCGATTAACAACAGTTGTCTCTGGAAACAGATGAATTCATTGAAAGTTAGTTCA	808		
QY	234 GCACCTCTAACAAGTTGGTTAAAGTTAAAGTACAGACAATCAACCGAACTCCCTGTGT	293		
Db	809 GCAGCTTAAACAACTTGTTGATGTGATGTCAGTGAAGTTCAAGATTAATCTCATTTGCTTTTGC	868		
QY	294 GGAAGAGGAATTAATGCTGTGTAAGTGAAGCAGCAGCCATGCCCTTAACGAGCGGAAGTAT	353		
Db	869 TGAAGAGAACTTAATGCTTATCAAGGTTTCTGTGAACATGCTCTGCGAGAGACATACT	928		
QY	354 TGAGCTTGCCCAAGTATTCGGGGGCCGCAATTGTGGAATCTCGAACAACCCGACCAT	413		
Db	929 AGAATATTCGAAATCTTCGGGGGAAATCTGTGAAGTTTCGATCACTGTACGTT	988		
QY	414 -CGAATGTTGGGGGACCCGGGTAAATAGTAGACATCCTCAGATGTTGG-CCAAGTTGG	471		
Db	989 ACAAGCTTACGGGGATCTCGACAAGAGGTTGCAATTCAAAAGGCTGTTGAGACTTATGG	1048		
QY	472 CATTTAAAGAGTGGCTCGAACCGGCAAAATTTGCTTTGTGTGCGGAATCCGCGCTCATAC	531		
Db	1049 CACTGTGAGGTGCGCAACAAGGCGAGTGTGCGCTGGTCGCGCAATCCGATGTGCAATTC	1108		
QY	532 GGAATATCT	540		

```
Db      1109 CAGGTACT 1117

RESULT 13
US-10-282-122A-15104
; Sequence 15104, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27/257,931
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15104
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Bordetella pertussis
US-10-282-122A-15104

Query Match      22.6%; Score 127.6; DB 13; Length 489;
Best Local Similarity 57.9%; Pred. No. 1e-32;
Matches 226; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Cy      49 ATGAACACACCCCTCTGTTTGTAGTAGAAGTAAGCGAGTCTAAACCCCATTCGCC 108
      1 ATGAACACGCGATTTCTGCTGATGAGAAACACACCGCGGCTGTGCGCGTGC 60
      109 GACATATTGGCCCGCGGTGTTTAAATGAGAGCTTGGCGGTGGCGGGAACAG 168
      61 GGCCTGTTCTCCGCGCGGCTACACATGAGACGCTGACCGTGGGCCACCGAGGAC 120
      169 GGGACGTTTCCCGCATCACATGTGTGCGCGGGGATGAGAACACCATCGAACACTG 228
      121 GCGACCGTGTGCGCGCTGACCGTGTGACCGTGTGCGACGAGTCAACAGATC 180
      229 ACCAAGCACTCTCAAGTTGGTTAAAGTAATTAACTAGACGACATCACCGAACTCCC 288
      181 ACCAAGACCTCTCAATCGCTGTGTGAGAGTGTCAAGTCTGAGACCTCAACGAGGCGG 240

Db      1109 CAGGTACT 1117

RESULT 14
US-10-282-122A-13405
; Sequence 13405, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13405
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Burkholderia fungorum
US-10-282-122A-13405

Query Match      22.4%; Score 126.6; DB 13; Length 489;
Best Local Similarity 57.8%; Pred. No. 2.2e-32;
Matches 225; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Cy      49 ATGAACACACCCCTCTGTTTGTAGTAGAAGTAAGCGAGTCTAAACCCCATTCGCC 108
      1 ATGAACACATCATTTCCGCTCTGTGAGAAACACACCGCGGCTTATCAAGCGTGGTT 60
      109 GACATATTGGCCCGCGGTGTTTAAATGAGAGCTTGGCGGTGGCGGGAACAG 168
      181 ACCAAGACCTCTCAATCGCTGTGTGAGAGTGTCAAGTCTGAGACCTCAACGAGGCGG 240

Db      1109 CAGGTACT 1117
```

Db 61 GGTCTTTTTCGGACCGGCTACACATTGAAACCTTAGCGTGGCTCCGACCGAAGAC 120
Qy 169 GGGAGCGTTTCCCGCATCAACATGTGTGCGGGGATGAGAACCATCGAACACTG 228
Db 121 CATTGCGTGTCCGCGATGACCATCGTCTCCATTGGGTGGAGCAGATGATCGAACAGATC 180
Qy 229 ACCAGCCACTCTACAAAGTGGTTAACTTAAGTACAGACATCACCGAACTCC 288
Db 181 ACCAAGCATCTGAACCGCTGATCGAGTGGTGAAGTGTGACCTTACCGAGGCGCC 240
Qy 289 TGTGTGAAAGGAAATTGATGCTGTGAAGTGAAGCCATGCCCTTACCGAGCGAA 348
Db 241 CACATGAGCGGAGGTGATGTTGATCAAGTGAAGGCGGTGCGAAGAACGTAGAG 300
Qy 349 GTGATTGACTAGCCCAAGTATTCGGGCGCCGATTTGGATATCTCCGAAGACCGCTC 408
Db 301 ATGAAGCGGATGTGATATTTTTCGGCGCCGATCATGACGTACCGAAAGACTTAC 360
Qy 409 ACCATGAAATGGTGGGGGACCGCGGTAA 437
Db 361 ACGATGAACTGACGGGCGGACGACAA 389

RESULT 15

US-10-282-122A-11953

; Sequence 11953, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 11953

LENGTH: 525

TYPE: DNA

ORGANISM: Burkholderia cepacia

US-10-282-122A-11953

Query Match 22.3%; Score 126; DB 13; Length 525;
Best Local Similarity 57.5%; Pred. No. 3,7e-32;
Matches 245; Conservative 0; Mismatches 180; Indels 1; Gaps 1;

Qy 48 CATGAACACACCTCTCTCTTTTAAAGTGAAGCCGAGTGTAAACCCGCACTTC 107
Db 36 CATGAACACATCATTTTCCCTCTCTGAGAAAGAAACCGGGCGCTGTGCGCGTGT 95
Qy 108 CGACATATTTTCCCGCGGTGTTTAACTTGAAGAGTGTGGCGGTGGCGGAGACA 167
Db 96 CGGCTGTCTTTTCCGCAACGGGCTAACAACATCGAAACCTTACGCTGGCGCCGACGA 155
Qy 168 GGGGAGCTTTCCCGCATCAACATGTTGTGCGGGGAGTGAACACCATCGAACACT 227
Db 156 TCAATCGCTGTGCGGCTCACACATCGTTCCATTGGCTCCGACGAGTGTACAGAGAT 215
Qy 228 GACCAAGCACTCTACAAAGTGTGTTAACATTAAGTACAGAGATCACCGAACTCC 287
Db 216 CACAGACATCTGAACCGCTTATCGAGTGTGAAGTGTGAACCTGACGAGCTGC 275
Qy 288 CTGTGTGAAAGGAAATTGATGCTGTGAAGTGAAGGCCAATGCCCTTACCGAGCGGA 347
Db 276 ACAATGGAACGGAGCTGATGCTGATCAAGTACGTCAGTGGGCAAGGAGCGCGAGA 335
Qy 348 AGTGAATTGACTAGCCCAAGTATTCGGGCGCCGATTTGGATATCTCCGAAGACCGGT 407
Db 336 AATGAAGCGGATGTGGAATTTTCGGGCGCCGATCATGACGTACCGAAGACCTTA 395
Qy 408 CACCATGAATGTGTGGGAGACCGG-GGTAAATGTAGCATCTCCAGATGTTGGCAA 466
Db 396 CACGATGAATTGACGGGCGCGAGCAAGAGCTGACGATTCATCCAGGGGCTGAGACG 455
Qy 467 GTTGGC 472
Db 456 GGGGCGC 461

Search completed: July 25, 2004, 13:28:33
Job time : 281.203 secs

Mon Jul 26 12:13:49 2004

us-09-893-033-17.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 03:59:07 ; Search time 1611.95 Seconds
(without alignments)

10466.900 Million cell updates/sec

Title: US-09-893-033-17

Perfect score: 565

Sequence: 1 gtggaatttaccacatg.....ccctgaatccaagttag 565

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_escba:*
2: em_escin:*
3: em_escin:*
4: em_escin:*
5: em_escin:*
6: em_escin:*
7: em_escin:*
8: em_escin:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_esc3:*
12: gb_esc3:*
13: gb_esc4:*
14: gb_esc5:*
15: em_escfun:*
16: em_escfun:*
17: em_escfun:*
18: em_escfun:*
19: em_escfun:*
20: em_escfun:*
21: em_escfun:*
22: em_escfun:*
23: em_escfun:*
24: em_escfun:*
25: em_escfun:*
26: em_escfun:*
27: em_escfun:*
28: gb_esc1:*
29: gb_esc2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	25.1	679	13	BQ407261 GA_Ed010
2	137	24.2	782	14	CB642034 OSJNEB01P
3	136	24.1	698	14	CB647745 OSJNEB01P
4	131.6	23.3	1380	11	AY105043 Zea mays

C	5	130.6	23.1	911	14	CA238460	SCRFPL503
C	6	129.8	23.0	799	14	CK291478	EST754192
C	7	129.8	23.0	962	14	CK297674	EST760388
C	8	129.4	22.9	867	14	CK243941	EST727578
C	9	129.4	22.8	584	14	CA242014	SCRFPL308
C	10	129.4	22.8	626	14	CB672821	OSJNE07B
C	11	129.4	22.8	695	14	BP184863	BP184863
C	12	128.8	22.8	697	14	CF210202	CAB20006
C	13	128.6	22.8	548	14	CD429148	ETH1.2.A0
C	14	128.6	22.8	782	14	BG581436	EST483170
C	15	128.2	22.7	891	14	CK243942	CK243942
C	16	128.2	22.7	938	14	CK257587	EST741234
C	17	128.2	22.7	972	14	CK243965	EST727602
C	18	128.2	22.7	977	14	CK255153	EST738790
C	19	127.8	22.6	664	14	BJ574968	BJ574968
C	20	127.8	22.6	711	12	BJ574939	BJ574939
C	21	127.6	22.6	541	10	BE995544	EST41267
C	22	127.4	22.5	728	14	CB978073	CB978073
C	23	126.6	22.4	885	14	CK258640	EST742277
C	24	126.4	22.4	638	10	BE823545	BE823545
C	25	125.8	22.3	684	14	CB654393	CB654393
C	26	125.6	22.2	710	14	CF397385	RTD3.3.B
C	27	124.6	22.1	703	14	CF480319	POL1.65.B
C	28	124.2	22.0	671	13	BU102997	SCFPL602
C	29	124.2	22.0	671	13	CA097418	CA097418
C	30	124.2	22.0	676	13	BU102998	SCFPL602
C	31	124.2	22.0	676	13	CA097419	CA097419
C	32	123.4	21.8	660	12	BP184425	BP184425
C	33	123.4	21.8	824	14	CB654392	CB654392
C	34	123.4	21.7	601	10	AM696137	AM696137
C	35	122.6	21.7	624	10	AM181183	687021B0.4
C	36	122.2	21.6	589	14	CD909456	CA68.112K
C	37	121.6	21.5	632	13	BQ415906	BQ415906
C	38	121.6	21.4	598	14	CD830820	BN40.0460
C	39	121.6	21.4	811	14	CK194674	FGAS00310
C	40	120.4	21.3	552	12	BJ310586	BJ310586
C	41	119.8	21.2	606	13	CA078461	SCRLAM100
C	42	119.2	21.1	797	12	BG836929	2M08.0940
C	43	118.8	21.0	702	12	BU169282	BU169282
C	44	118.8	21.0	702	12	BU581766	BU581766
C	45	118.6	21.0	563	13	CA021505	H240B18r

ALIGNMENTS

RESULT 1
BQ407261
LOCUS GA_Ed0104F04f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION GA_Ed0104F04f Gossypium arboreum 7-10 dpa fiber library Gossypium
arborescens cDNA clone GA_Ed0104F04f, mRNA sequence.

ACCESSION BQ407261
VERSION BQ407261.1 GI:21094948
KEYWORDS
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum

REFERENCE
Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber

TITLE
Unpublished (2000)
JOURNAL
COMMENT
Contact: Wing R.
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Total High Quality bases = 441

Seq primer: TAAATGAGTCACTATAGGG
High quality sequence start: 6
High quality sequence stop: 620.
Location/Qualifiers

FEATURES

1..679
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0104P04f"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 25.1%; Score 142; DB 13; Length 679;
Best Local Similarity 57.0%; Pred. No. 2.5e-29;
Matches 259; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 55 CACACCTCTCTGTTTAAAGTGAAGATGAGCCGAGTGTAAACCCGATTCGGACTA 114
DB CACTTATTCATCTGTGTAATGACTCTCCGGAGTTCTAAACCTGTACAGGTGT 285
QY 115 TTGCCCCCGCGTTTAAACATTGAGCTTGGCGGTGGGCTGGCGGACAGGGGAC 174
DB TTGCTCGAAGGCTAATCTTCAGATTGGCTGTGGAACAGCTGAAGTTAGGGG 345
QY 175 GTTTCGCCATCACCATGATGATGCGGGGAGTGAAGACACCATCGAACAATGACCAAG 234
DB CTTTCCGATTAACAACGTGTGCTCCGTGTAAGTGAATCAATTAGCAATTTAGTCAG 405
QY 235 CAACTCTAAGCTGTGTAAGTGAATTAAGTACAGGACATCACCAGAACTCCCTGTGTG 294
DB CAACCTAATGAGCTGTGATATGATAGATGAGTACGGGATCTTACACAGTTGCCATTTGCT 465
QY 295 GAAAGGAAATGATCTGTGATGAGTGAAGTGAAGCCCAATGCCCTTAACCGAGCGAAGTATT 354
DB GAACAGAACTAATTTATTAAGTATGCTGTGAATGCACTGCTGACAGAGTGTCTT 525
QY 355 GAGCTAGCCCAAGTATTCGGGCGCCGCAATTGTGATATCTCCGAAGACACCGTACCATC 414
DB GACATTTGCCAATTTTAAAGGCGCCAAAGCTGTGATGTTTTCGACACACCGTCAACCTC 585
QY 415 GAATGATGGGGGACCCGGGTAAATGTGAGCAATCTCCAGATGTGGCCAAAGTTGCGAT 474
DB GAGCTACAGAGATTTAACAATAATGTTGCACTGCAACAAAGATTGTTGAACCCACCGTAT 645
QY 475 TAAAGAGGTGCTCGAACGGCCAAATTTGCTTTG 508
DB CTGTGAAGTTGCTAAACTGGAACCTCGCACTG 679

RESULT 2
CB642034/c 782 bp mRNA linear EST 08-APR-2003
LOCUS CB642034
DEFINITION OSJUNBD1P24.r OSJUNBD Oryza sativa (japonica cultivar-group) cDNA
ACCESSION CB642034
VERSION CB642034.1 GI:29637025
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 782)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL

Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3867
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: p column: 24
Seq primer: gga aac agc tat gac cat g.

FEATURES

Location/Qualifiers

ORIGIN

Query Match 24.2%; Score 137; DB 14; Length 782;
Best Local Similarity 57.7%; Pred. No. 7.1e-28;
Matches 282; Conservative 0; Mismatches 205; Indels 2; Gaps 2;

QY 54 AACACCCCTCTCTGTTTAAAGTGAAGATGAGCCGAGTGTAAACCCGATTCGGACTA 113
DB AATTTCTTTGTCATTTTGTGTAAGATTTCCCTGAGTCTCAATGTTTGAACGGTGT 640
QY 114 ATTGCCCCCGCGTTTAAACATTGAGCTTTGCGGTGGGCTGGCGGAAACAGGGGGA 173
DB TTTCCTCCGAAAGGCTAATATTCAGAGTGTGCTGTGGCCAGCTGAAAAAGAG 580
QY 174 CTTTCCGCAATCACCATGATGATGCGGGGATGAAACACATCGAACAATGACCA 233
DB CACTTCTCGATCACTACTGTGTGCTCCGGAACATGATGATTAATGCAAGCTAGTACA 520
QY 234 GCAACTCTCAAGTGTGTTAAGTAAATTAAGTACAGACATCAACCGAACTCCCTGTGT 293
DB CCAACTGTAAAGCTCATGATGATTTATGAGGTCCAGATCTTACTATTACATTTGC 460
QY 294 GGAAGGAAATGATGCTGTGAGAGTGAAGCGCCAAATGCCCTTAACGAGCGAAGTAT 353
DB TCTTAGAGAGTAAAGATCATTAAGATTGCTGTAAACACACAGCCCGCAGGGCTATCT 400
QY 354 TGAGCTAGCCCAAGTATTCGGGCGCCGATTTGATATCTCCGAAGACACCGTAC-CA 412
DB AGATTTGCTGATATTTTCCGGCCAAACGTGTGATGATATCAACCGTAACTCT 340
QY 413 TCGAATGTTGGGGAACCCGGTAAATGTGTAACATCTCCAGATG-TTGGCCAAAGTTGG 471
DB TCAGCTTACTGAGACCTTGATTAATAGTGTGCAATTACAAAGAGATGCTCGAGCCCTATGG 280
QY 472 CATTAAGAGTGTGCTCGAAGCGGCAAAATGCTTTGTGTGGGGAATCCGGCTCAATAC 531
DB CATTTGTAGGTTGACGAACCTGGAAGGAGGTTGCTGTGGCGGTGAGTCAAGAGTGCATTC 220
QY 532 GGAATATCT 540
DB 219 CAATATCT 211

RESULT 3
CB647745/c

LOCUS	OSUNEB1.0M22	698 bp	mRNA	linear	EST 08-APR-2003
DEFINITION	OSUNEB1.0M22.r OSUNEB Oryza sativa (japonica cultivar-group) cDNA				
ACCESSION	CB647745				
VERSION	CB647745.1				
KEYWORDS	GI:29642738				
SOURCE	EST.				
ORGANISM	Oryza sativa (japonica cultivar-group)				
REFERENCE	Oryza sativa (japonica cultivar-group)				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
TITLE	1 (bases 1 to 698)				
JOURNAL	Jantanasurawat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.				
COMMENT	Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe oryzae				
	Unpublished (2003)				
	Contact: Rod Wing				
	Arizona Genomics Institute				
	University of Arizona				
	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA				
	Tel: 520 626 3967				
	Fax: 520 621 9288				
	Email: http://genome.arizona.edu				
	PCR Primers				
	FORWARD: gta aac cga cgg cca gtc				
	BACKWARD: gga aac agc tat gac cat g				
	Plate: 10 row: M column: 22				
	Seq primer: gga aac agc tat gac cat g.				
FEATURES	Location/Qualifiers				
source	1..698				
	/organism="Oryza sativa (japonica cultivar-group)"				
	/mol_type="mRNA"				
	/cultivar="Nipponbare"				
	/db_xref="taxon:39947"				
	/clone="OSUNEB1.0M22"				
	/tissue_type="leaf"				
	/dev_stage="3 week"				
	/lab_host="DH10B"				
	/clone_lib="OSUNEB"				
	/note="vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (Cne 86061)."				
ORIGIN					
Query Match	24.1%; Score 136; DB 14; Length 698;				
Best Local Similarity	57.6%; Pred. No. 1.3e-27;				
Matches 261; Conservative 0; Mismatches 205; Indels 2; Gaps 2					
OY	55	CACACCCCTCTCTGTTT	AGTTGAAGTGAAGCCGAGTCTAACCCGACTTGC	CGGACTA	114
DB	698	CATTCTTTGTCATTGTTGGTGAATGATTTCCCTGGAGTTC	CAATGTTGAACAGTGTT		639
OY	115	TTTGGCCCGCGTGT	TTTTTAACTTGAAGAGCTTGGCGTGGGTGCGCGGAA	CAGGGAGAC	174
DB	638	TTCTCCCGAAGGCTACA	TAATTCAGAGTCTGCTGTGGCCCACTGMAA	AAAAAAGGC	579
OY	175	GTTTCCCGCATC	CAATGAGTGTGTCGCGGGGAGTGAACACCATGAA	CAACTGACCAAG	234
DB	578	ACTTCTCGGATAC	TACTGTTTCCCTGGAGTGAATGATCTATTG	CCAAAGCTAGTACAC	519
OY	235	CAACTCTCAAGTTG	GTGTTAAGCTAATTAAGTACAGGACATACCGAA	AACTCCCTGTGG	294
DB	518	CAACTGTATTA	CTCAATGATTTTATGAGTGCCAGATCTTACTCATTTT	TACCAATTTTGT	459
OY	235	GAAAGGGAATTG	ATGCTGTTGAAGTGAGCGGCAATGCCCC	TAAACCGAGCGGAAGTGATT	354
DB	458	GCTAAGAGATTAT	ATATCATTAAGAATGCTGTAAACACACAGCCCGCAGGCGCTATCTCA		399
OY	355	GAGCTAGCCCA	GTGTTTCCGGGCGCCGCAATGTGGATATCTCCGAGAGAC	CCGCTCAC-CAT	413
DB	398	GATATGCTGAT	ATATTTTCCGGGCGCAAACTGTGGAGTATCAAGATCACACCGTAACTCTT		339

Oy	414	CGAATGTGGGGAGCCGGGTAATAATGTTGCAATCTCCAGATG--TTGGCCAAAGTTGGC	472
Db	338	CAGGTTACTGAGAACCTTGATTAATAATGTCGCAATTACAAAGATGCTCGAGCCCTAAGCC	219
Oy	473	ATTAAAGAGGTGCTCGAACGGCGCAAAATTCCTTTGTGTCGGGAAATCCGGCGTCAATACG	532
Db	278	ATTGTAGAGGTGACCAATCTGGACGAGGTTCGTTGCGCCGTGAGTACGAGATTCGATTCC	219
Oy	533	GAATATCT 540	
Db	218	AAATACCT 211	
RESULT 4			
LOCUS	AY105043	1380 bp	mRNA linear HTC 16-OCT-2002
DEFINITION	Zea mays FC0069007 mRNA sequence.		
ACCESSION	AY105043		
VERSION	AY105043.1	GI:21208121	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 1380) Hainey,C.F., Dolan,M., Mao,G.H., Vogel,J.M., Wuitsilt,M.S., Arthur,L.W., Hanafey,M., Morange,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 1380) Coe,E.H.		
AUTHORS	Direct Submission		
TITLE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
JOURNAL			
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZMDb and may be found by BLAST searching at MSL, maizemap.org; ZMDb, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZMDb: www.zmdb.iastate.edu.		
FEATURES	Location/Qualifiers		
source	1..1380		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/db_xref="MaizeDB:633742"		
	/db_xref="taxon:4577"		
	/clone_id="Maize Mapping Project/Dupont Consensus Library"		
	/note="this sequence is part of a project of EST assemblies resulting from the application of EST contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
ORIGIN			
Query Match	23.3%;	Score 131.6;	DB 11; Length 1380;
Best Local Similarity	57.3%;	Pred. No. 3.3e-26;	
Matches	276;	Conservative 0;	Mismatches 204; Indels 2; Gaps 2;
Oy	61	CTCTCTGTTTAGTTGAAGATGAAGCCGAGATGCTAACCCGCATTCGCGACTATTTGGCC	120
Db	635	CTCTCCATCCTTGGAAATGACTGCTCGGTGCTCTCAACATTGTAACAGGAGTCTTTGCT	694
Oy	121	CGCGGTGTTTAACTTGAGAGCTTGCGGTGGGCTCGCGCGAACAAGGGGAGAGTTCC	180
Db	695	CGCAGGGGCTACATATACAGAGCTTGCTGTTGGCCCACTGAGAGGAAGGAGATTTCG	754
Oy	181	CGCATCAACATGATGTGTGCGCGGGAGTGAACAACCATCGAACACTGACCAACCAACTC	240

Db 755 CGGATTACACAGTGTCTCTGTAAGTCAATGAGAGTTAGTTCACACCTT 814
 QY 241 TACAAAGTGTAACTGTAATTAAGTACAGACATACCGAACTCCCTGTGTGAAAG 300
 Db 815 TACAAAGTGTAACTGTAATTAAGTACAGACATACCGAACTCCCTGTGTGAAAG 874
 QY 301 GAATGATGCTGTGTAAGTGTAGCGCAATGCCCTTACCGAGCGGAAGTATGAGCTA 360
 Db 875 GAATGATGCTGTGTAAGTGTAGCGCAATGCCCTTACCGAGCGGAAGTATGAGCTA 934
 QY 361 GCCAGGATTCGGGGCCCGCATTTGTGATATCTCCGAAGACACCGTCAC-CATCGAATG 419
 Db 935 GGTGAATCTTCGCGCAAAACCTGTGATGTTTCTGACACACAGTAAACGCTTCACTT 994
 QY 420 GTGGGGGACCCGGGTAATGTGTGCAATCTCCAGATGTTGG-CCAAAGTTGGATTA 478
 Db 995 ACTGGAATCTTGAAGAGATGTTGCACTACAAAGTTATTTAGACCTATATGATCTGT 1054
 QY 479 GAGTGAGTGTGAAACGGGCAAAATGCTTTGTGTGCGGGAATCCGGCTCAATACGAAATAT 538
 Db 1055 GAGTGCGCAAGTGTGACAGATGCACTGTGTCGGCAATCGAAGTGTGACTCCAGTAC 1114
 QY 539 CT 540
 Db 1115 CT 1116

RESULT 5
 CA238460/c 911 bp mRNA linear EST 25-SEP-2003
 LOCUS SCRFPL5036A10.g Saccharum officinarum Flus Saccharum officinarum
 DEFINITION CDNA clone SCRFPL5036A10 5', mRNA sequence.
 ACCESSION CA238460
 VERSION CA238460.1 GI:35310054
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum.
 1 (bases 1 to 911)
 Vector: A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parnuda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 036 row: A column: 10
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1..911
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCRFPL5036A10"
 /lab_host="DH10B"
 /clone_lib="Saccharum officinarum FL5"
 /note="Organ: Developed inflorescence (20cm-long) without
 rachis; Vector: pSPORT1; Site_1: SalI; Site_2: NotI; An
 unidirectional cDNA library generated from [Developed
 inflorescence (20cm-long) without rachis]. cDNA was
 prepared from polyA+ mRNA using Superscript Plasmid
 System Kit (Invitrogen). The double-strand cDNAs were
 fractionated in a sepharose CL-2B 40cm-columns and
 fragments sizing between 0.8 and 1.5 Kb were

ORIGIN
 Query Match 23.1%; Score 130.6; DB 14; Length 911;
 Best Local Similarity 56.9%; Pred. No. 5, 2e-26;
 Matches 275; Conservative 0; Mismatches 206; Indels 2; Gaps 2;

directionally cloned into the vector. Details of each
 source of RNA and library construction can be obtained at
 http://succest.lad.ic.unicamp.br/public"

Db 60 CCTCTGTGTTTAAAGTGAAGTGAAGCCGAGTGTAAACCGCATTTGCCGACTATTTC 119
 QY 646 CCTCTCATCTCTTGAATGAGCTGTCNTGTGTCTTCAACATTTGAACAGAGTTCTTTC 587
 Db 120 CCGCCGTGTTTAAACATTTAGAGCTTGGCGGTGTGCGGGAACAGGGAGCTTTC 179
 QY 586 TGGCAGAGGTGCAATATATACAGACCTTCTGTGGCCAGCTGAGGAAGGCAATTC 527
 Db 180 CCGCATCACATGTGTGTGTCGGGGAGTGAACACCATGCAACACTGACCAACACT 239
 QY 526 GGGTATTACAAAGTGTTCCTGTGATCTGTGAATCCATTGGAAGTTAGTTCAACACT 467
 Db 240 CTACAAAGTGTGTTAAGTAAATTAAGTACAGAGCATCCAGAACTCCCTGTGTGAAG 299
 QY 466 TTACAGCTGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 407
 Db 300 GGAATGATCTGTGTAAGTGAAGTGAAGCGCAATGCCCTTAACCGAGCGAATGATGACT 359
 QY 406 GGAATGATCTGTGTAAGTGAAGTGAAGCGCAATGCCCTTAACCGAGCGAATGATGACT 347
 Db 360 AGCCAGGATTCGGGGCCCGCATTTGTGATATCTCCGAAGACACCGTCAC-CATCGAAT 418
 QY 346 TGCCGAATCTTCGAGCAAAACCTGTGATTTTCTGACCTTAAGTGAAGTGAAGCTTCACT 287
 Db 419 GGTGGGGAACCGGGTAATGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 477
 QY 286 TACTGAGATTTTGAACAGATGTTGCACTACAAAGTTATTTGAGGCAATGATGATCTG 227
 Db 478 AGAGTGTGTGCAACGGGCAAAATGCTTTGTGTGGGAATCCGCGCATACGAAAT 537
 QY 226 CAGGTGCGCAAGTGTGACGAGTGTGCTGTGTAATGGAAGTGTGACTCCAAATGA 167
 Db 538 TCT 540
 QY 166 CCT 164

RESULT 6
 CK291478 799 bp mRNA linear EST 15-DEC-2003
 LOCUS EST754192 Nicotiana benthamiana mixed tissue cDNA library,
 DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NBNC250 5',
 end mRNA sequence.
 ACCESSION CK291478
 VERSION CK291478.1 GI:39871957
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots;
 asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
 1 (bases 1 to 799)
 Buell, C.R., Hart, A., Zismann, V., Kazemiyeva, S.A., Day, B.,
 Staskawicz, B., Jin, H. and Baker, B.
 Generation of EST sequences from Nicotiana benthamiana
 Unpublished (2003)
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.
 Location/Qualifiers
 1..799

```

/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBYC250"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/notes="Vector: pCMVSPORT6.1; Site_1: EcoRI, Site_2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

ORIGIN

```

Query Match      23.0%; Score 129.8; DB 14; Length 799;
Best Local Similarity 58.1%; Pred. No. 8,2e-26;
Matches 270; Conservative 0; Mismatches 187; Indels 8; Gaps 2;

55 CACACCTCTCTGTTTGTAGTGAAGTGAAGCCGAGTCTAACCCGATTGCCGACTA 114
280 CACACCATTTTCAGTATTTGTTGGGATGAAGTGTATATTAACAGAAATTGCTGAGTT 339
115 TTTCGCCGCGTGTGTTTAACTTGAAGCTTGGCGTGGGCTGCCGGAACAGGGGAC 174
340 TTTCGTCGAAGGGGTTAACAATTAATGACCTGCTGTGGTTGACAA-----GGAT 393
175 GTTTCGCCGATCACCATTGTTGTCGCGGGGATGAACACATCGAACACTGACCAAG 234
394 AAAGCTCTCTTACCAATGTTCTCTGGAACAGAAAGCTTGCACAAAGTTGTAGAG 453
235 CAACCTCAACAAGTTGTTAAGCTTAATTAAGTACAGACATCACCGAAATCTCCTGTG 294
454 CAGCTCAACAAGCTTGTGAAGTCTTGAAGTGAAGGATCTGTCAAGAGAACCAAGTT 513
295 GAAAGGGAATTGATCTGTGTGAAGTGAAGCCGCAATGCCCTAACCGCGGAAGTGT 354
514 GAACGAGAAATGATCTCAATAAGCTGAATGACATGACAGACAAAGCTGAATTTAG 573
355 GAGTAGCCCAAGTATTCGCGGCCCGCATTTGATGATCTCGAAGACACGTCACATC 414
574 TGTTGTGTGACATCTTCAGAGGAAATTTGTGATATATACAGACATTTTGTGACTAG 633
415 GAATGG--TGGGGACCCCGGTAAATGTGTAGCAATCTCCAGATTTGGCCAACTTGGC 472
634 GAGGTGACTGTGACCCAGGAAATGTGCTGTGTTTGAAGAAATTTAAGCAAGTTGGG 693
473 ATTAAGAAGTGTGCTGAACGGGCAAAATTTGCTTTGATGCGGGA 517
694 ATCAAGAAGTACCCGAACTGGAAGATGCTTTGAGACGTGA 738

```

RESULT 7

```

CK297674      962 bp      mRNA      linear      EST 15-DEC-2003
LOCUS         CK297674
DEFINITION    ESTT60388 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMDS4 5',
end, mRNA sequence.

```

```

ACCESSION    CK297674
VERSION      CK297674.1 GI:39884298
KEYWORDS     EST.
SOURCE       Nicotiana benthamiana
ORGANISM     Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

```

```

REFERENCE     1 (bases 1 to 962)
AUTHORS      Bell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
TITLE        Generation of EST sequences from Nicotiana benthamiana
JOURNAL      Unpublished (2003)
COMMENT      Other ESTs: EST760399
Contact: Robin Bueli
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

```

FEATURES

source

```

1..962
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBMDS4"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/notes="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

ORIGIN

```

Query Match      23.0%; Score 129.8; DB 14; Length 962;
Best Local Similarity 58.1%; Pred. No. 9e-26;
Matches 270; Conservative 0; Mismatches 187; Indels 8; Gaps 2;

55 CACACCTCTCTGTTTGTAGTGAAGTGAAGCCGAGTCTAACCCGATTGCCGACTA 114
291 CACACCATTTTCAGTATTTGTTGGGATGAAGTGTATATTAACAGAAATTGCTGAGTT 350
115 TTTCGCCGCGTGTGTTTAACTTGAAGCTTGGCGTGGGCTGCCGGAACAGGGGAC 174
351 TTTCGTCGAAGGGGTTAACAATTAATGACCTGCTGTGGTTGACAA-----GGAT 404
175 GTTTCGCCGATCACCATTGTTGTCGCGGGGATGAACACATCGAACACTGACCAAG 234
405 AAAGCTCTCTTACCAATGTTCTCTGGAACAGAAAGCTTGCACAAAGTTGTAGAG 464
235 CAACCTCAACAAGTTGTTAAGCTTAATTAAGTACAGACATCACCGAAATCTCCTGTG 294
465 CAGCTCAACAAGCTTGTGAAGTCTTGAAGTGAAGATCTGCAAGAGAACCAAGTT 524
295 GAAAGGGAATTGATGCTGTGAAGTGAAGCCGCAATGCCCTAACCGAGGGAAGTAT 354
525 GAACGAAATTTGATGCTCATTAAGCTGAATGCAAGTGAAGCAAAAGCTGAATTTAG 584
355 GAGTAGCCCAAGTATTCGCGGCCCGCATTTGATGATCTCGAAGACACCGTACCATC 414
585 TGTTGTGTGACATCTTCAGAGGAAATTTGTGATATATACAGACATTTTGTGACTATC 644
415 GAATGG--TGGGGACCCCGGTAAATGTGTAGCAATCTCCAGATTTGGCCAACTTGGC 472
645 GAGGTGACTGTGACCCAGGAAATGTGCTGTGTTTGAAGAAATTTAAGCAAGTTGGG 704
473 ATTAAGAAGTGTGCTGAACGGGCAAAATTTGCTTTGATGCGGGA 517
705 ATCAAGAAGTACCCGAACTGGAAGATGCTTTTGAAGAGTGA 749

```

```

RESULT 8
CK243941/c 887 bp mRNA linear EST 12-DEC-2003
LOCUS EST127578 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone F0CA635 3' end, mRNA sequence.
ACCESSION CK243941
VERSION CK243941.1 GI:39789035
KEYWORDS EST
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 887)
REFERENCE Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A. and Baker, B.
AUTHORS Generation of ESTs from potato callus tissue
TITLE Unpublished (2003)
JOURNAL Other ESTs: EST127579
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-arr@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT.
Location/Qualifiers
1..887
/organism="Solanum tuberosum"
/mol_type="mRNA"
/db_xref="taxon:4113"
/clone="F0CA635"
/tissue_type="callus"
/lab_host="DH10B-TomA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/notes="Vector: pGMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."
ORIGIN
Query Match 22.9%; Score 129.4; DB 14; Length 887;
Best Local Similarity 55.4%; Pred. No. 1.1e-25;
Matches 271; Conservative 0; Mismatches 216; Indels 2; Gaps 1;

QY 54 ACAACCTCTCTCTGTTTGAAGTGAAGCCGAGTGTACCCGCAATGGCCGACT 113
DB 809 ACAACATTAAAGCATGCTGTGCAACAAATGCTCCGAGGCTCTTAATTGATGCTGT 750
QY 114 ATTGCCCGCGCGTGTATTACATTGAGAGCTTGGCGGTGGCGGAGACAGGCGGA 173
DB 749 TATATCTGAGAGAGGATTAATTTTCAGAGCTTCTGTAGGCTTCAAGACCAAGG 690
QY 174 CGTTTCCGATCAACATGCTGTGCGGGGAGATGACACCAATGCAACTGACCA 233
DB 689 ACTTTCAGTATTTACCAAGTGTCTCTGGAATGATGAATAATAGGAAATTGTCAA 630
QY 234 GCAACTCTACAGTGTGTTAACGTAATTAAAGTACAGACATACCGAAATCCCTGTGT 293
DB 629 GCAATTTAACAAGTTGATGATGTTCAATGAGTTCAAGATTAATCTCATTACCATTTGG 570
QY 294 GAAAGGGAATGATGCTGTGAGAGTGAAGCGCCATGCCCTTACCGAGGAGATGAT 353
DB 569 TGACCGGAGCTATGCTCATTAAGTAGCCGCAATGACAGCAAGAGGAGATGCTT 510
QY 354 TGACTTACCCAGTATTTCCGGGCGCATTTGAGATATCTCCGAGACACCGTCACT 413
DB 509 TGATATTCACAGTATTTTCCGTCTAAACCTGTGATGTGATCATCACTAATCACTT 450
QY 414 CGAATGCTGGGGGACCCGGGTAATATGTAGCAATCTCCAGAT--GTTGGCCAATGG 471
DB 449 GGAATTAAGTGAAGACTTCAACAAGCTGTGGCATTCACAAAGTATGAGGCTTATGG 390
QY 472 CATTAAGAGGTGCTGAAAGCGCAAAATTTGCTTGTGTCGGAAATCCGGCTCAATAC 531

```

```

DB 389 CATTGTGAGGTGCTCTGTAACAGGCGGATGACATTGTGACAGAAATCTGTGTGACTC 330
QY 532 GGAATATCT 540
DB 329 GACATACCT 321

RESULT 9
CA242014
LOCUS SCEPFL3082B07.g Saccharum officinarum FL3 Saccharum officinarum
DEFINITION cDNA clone SCEPFL3082B07 5', mRNA sequence.
ACCESSION CA242014
VERSION CA242014.1 GI:35317200
KEYWORDS EST
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 584)
REFERENCE Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
AUTHORS The libraries that made SUCEST
TITLE Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
JOURNAL Contact: Arruda P
COMMENT Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parnuda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 082 row: B column: 07
Seq primer: 17 Promoter Primer.
Location/Qualifiers
1..584
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEPFL3082B07"
/lab_host="DH10B"
/clone_lib="Saccharum officinarum FL3"
/notes="Organ: Base of developing inflorescence (5cm-long);
Vector: pSPORT1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from (base of
developing inflorescence (5cm-long)). cDNA was prepared
from polyA+ mRNA using Superscript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://succest.iad.ic.unicamp.br/public"
ORIGIN
Query Match 22.8%; Score 129; DB 14; Length 584;
Best Local Similarity 58.0%; Pred. No. 1.2e-25;
Matches 265; Conservative 0; Mismatches 190; Indels 2; Gaps 2;

QY 86 CCGAGTCTAACCCGATTCGCCGACATTTTGGCCCCCGGTGTTTAACATTGAGACT 145
DB 9 CTGGGTCTCAACATTGTAAACAGAGCTTTGCTCGAGGGGCTAACATATACAGAGCC 68
QY 146 TGGCGTGGGGTTCGGCGGACAGGGGAGACGTTTCCGCCATCACCATGATGAGTGGCGGGG 205
DB 69 TTGGTGTGGCCAGCTGAGAGAGGACATTTGGCGGTATTAACAAGTGTCTCTGTA 128
QY 206 ATGAGAACACCATGAGACAACTCTTCAACAGTCTTGAAGTTAAAG 265
DB 129 CTGTTGATTCATTGAGAGATTAGTTCAGCAGCTTTTCAAGCTTGTGATGTCATGAG 188

```

OY 266 TACAGACATCACCGAAACTCCCTGTGTGAAAAGGAAATTGATGCTGTGTAAGGTGACG 325
DB 189 TTGATGACATTAACCCCTCACTTTTGTGTGAAAAGGAACTGATCTATTAAAGTTTCTG 248
OY 326 CCAATGCCCTTAACCGAGCGGAAGTGAATTGACCTAGCCCAAGTATTTCCGGCCCGCATTTG 385
DB 249 TAAACACTGCTGCTGCGAGGAAATCTAGATATTGCGGAATCTTCCGACAAACCTG 308
OY 386 TGGATATCTCCGAGAGACACCGTAC -CATGGAATGTGCGGGGACCCGGGTAAATGTAG 444
DB 309 TTGATGTTTCTGACCATTAAGTAAAGCTTCAAGTACTTACGAGATTTTGACAAATGTTG 368
OY 445 CAATCTCCAGATGTTGG -CCAAGTTGCAATTAAGAGGTGCTCGAACCAGGCAAAATTG 503
DB 369 CACTACAAAGGTATTGAGGACCATATGCGCATCTCGAGGTGCGCAAACTCGACGAGTGG 428
OY 504 CTTTGTCGGGGAATCCGCGTCAATACGGAATATCT 540
DB 429 CACTGTCTGTAATCGAAGGTGCACTCCAAATCACT 465

RESULT 10

CB672821/c

LOCUS CB672821 626 bp mRNA linear EST 09-APR-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA
clone OSUNEB07B04.3, mRNA sequence.

ACCESSION CB672821
VERSION CB672821
KEYWORDS GI:29676546
SOURCE EST.

ORGANISM

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 626)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,D., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.

TITLE

Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL

Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA

COMMENT

85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers

FEATURES

source
1..626
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEB07B04"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_1lb="OSUNBE"
/note="Vector: pBluescript II KS +; site 1: EcoRI; site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

ORIGIN

Query Match 22.8%; Score 129; DB 14; Length 626;
Best Local Similarity 56.6%; Pred. No. 1.2e-25;
Matches 277; Conservative 0; Mismatches 210; Indels 2; Gaps 2;

OY 54 ACACACCTCTCTGTTTATTGTTAAAGTAGAACCCGAGTGTCTAACCCGATTGCCGACT 113
DB 623 ACATATCTATTCATCTCTGTGTAATGATTTGCCCTGTGTTTCAACATTGTTACAGGGGT 564
OY 114 ATTTGCCGCGGNGGTTTAAACATTGAGAGCTTGGCGGTGGGCGGACAGAGGGA 173
DB 563 CTTTGCTGCAAGGCTACATATACAGAGCTTGTGTAGGCCACGCTGAAAGTCAAG 504
OY 174 CGTTTCCGCAATACCATGCTGTGTCGCGGGGATGAGAACACCATGAAACAATGACCA 233
DB 503 CTTTGCGGTATTACCAACAGTTGCTCTGGAACAAGATGATCATTTGAAATGATTGA 444
OY 234 GCAACTCTCAAGTTGTTAAGTAAATTAAGTACAGGACATCACGAAATCCCTGTGT 293
DB 443 GCACTTAAACAATCTGTGATGTGATGATGATGATGATGATGATGATGATGATGAT 384
OY 294 GGAAGGAAATGATGCTGTGAGAGTGAACGCAATGCCCTTAACCGAGCGAAGTAT 353
DB 383 TGAAGAGAACTATGCTTATCAAGGTTCTGTGAACAATGCTGCTCGAGAGACATCT 324
OY 354 TGAAGTACCGCAGATATTTCCGGCCCGCATTTGTGATATCTCCGAAGACACCGTCA 413
DB 323 AGATATGCTGAATCTTCCGGGCAAAATCTGTGATGTTTGTGATGATGATGATGATG 264
OY 414 -CGAATGTGTGGGAGACCCGGGTAAATGTGTAATCTCTCAGATGTTGG -CCAAGTTG 471
DB 263 ACAGCTTACTGGGATCTGACAAAGATGTTGATTAACAAGCTGTTGAGAGCTTATG 204
OY 472 CATTAAAGAGGTGCTCGAACCAGGCAAAATGCTTTGTCGCGGAATCCGGCTCAATC 531
DB 203 CACTGTGAGGTGCGCAAGACAGGGCGAGTGGGCTGTGCGGATCCGATCCGATTC 144
OY 532 GGAATATCT 540
DB 143 CAAGTCACT 135

RESULT 11

BP184863

LOCUS BP184863 695 bp mRNA linear EST 17-JUL-2003
DEFINITION BP184863 pms rice callus cDNA Oryza sativa (japonica
cultivar-group) cDNA, mRNA sequence.

ACCESSION

BP184863
BP184863
GI:32948291

VERSION

EST.
Oryza sativa (japonica cultivar-group)

KEYWORDS

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 695)
Moriguchi,K., Ito,Y., Yamazaki,Y. and Kurata,N.

AUTHORS

Findings of various plant nuclear proteins using yeast nuclear
transportation trap system - a proteomal approach

TITLE

Unpublished (2003)
Contact: Kazuki Moriguchi

JOURNAL

Plant Genetics
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411-8540, Japan
Tel: 81-55-981-6872
Fax: 81-55-981-6879

COMMENT

Email: kmoriguc@lab.nig.ac.jp
cDNA clone obtained from nuclear transportation trap system
encoding a protein similar to Arabidopsis thaliana acetolactate
synthase-like protein.
location/Qualifiers

FEATURES

source
1..695

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/tissue_type="callus"
/dev_stage="4 days after regeneration treatment"
/clone_1lb="pMS rice callus cDNA"

ORIGIN

Query Match 22.8%; Score 129; DB 12; Length 695;
 Best Local Similarity 56.6%; Pred. No. 1.3e-25;
 Matches 277; Conservative 0; Mismatches 210; Indels 2; Gaps 2;

54 ACACACCCCTCTCTGTTTAAAGATGAAGCCGGAATGCTAACCCGCAATTCGCGACT 113
 |||
 62 ACATCTCTATCCATCTCTGTCATGATGCTGCTGTTTCAACATTTGTACAGGGGT 121
 |||
 114 ATTTCGCCCGGCTGTTTAAATGAGAGCTTGGCGGTGGGGTGGGGAACAGGGGA 173
 |||
 122 CTTTCTGCGCAGAGGCTACATATACAGAGCTTGTGTAGCCCACTGAAAGTACAG 181
 |||
 174 CGTTTCCCGCATACCAATGCTGTGCGCGGAGTGAACACCATTCGAACATGACCAA 233
 |||
 182 CTTTCGGGATTTACACAGTGTCTCCGGAACAGATGATCCATGAGAAGTATGCA 241
 |||
 234 GCAACTCTACAGTGTGTTAAGTAAATTAAGTACAGACATCCGAACTCCCTGTGT 253
 |||
 242 GAGGTTAACAACTGTTGATGTGATGATGATGATGATGATGATGATGATGATGAT 301
 |||
 294 GGAAGGGAATTGATGCTGTGTAAGTGAAGGCGCAATGCCCCCTAACCGAGGAAATGAT 353
 |||
 302 TGAAGAGAACTTATGCTTATCAAGGTTTCTGTGAACATGCTGCTCGAGAGACATACT 361
 |||
 354 TGAGTACGCCAGGATTTCCGGGCGCCGATGTTGGATATCTCCGAACACCGTCAACAT 413
 |||
 362 AGATATTCCTGAATCTTCGCGGCAAAATCTGTGATGTTCTGATCACCTGTACGTT 421
 |||
 414 -CGAATGTGGGAGACCCGGGTAATGTAAGCAATCCCATGATTTGG-CCAAATTTGG 471
 |||
 422 ACAGCTTACTGGGATCTCGCAAGATGTTGATTAACAAAGGCTGTGAGCTTATGG 481
 |||
 472 CATTAAGAGGTGCTGGAACGGGCAAAATTTGTTGGTGGGGAATCCGGCTCAATAC 531
 |||
 482 CATGTGAGGTCCCGCAAGAGGCGAGTGGCGCTGTCGCGAAATCCGATGTCATTC 541
 |||
 532 GGAATATCT 540
 |||
 542 CAAGTACCT 550
 |||

RESULT 12
 CP210202
 LOCUS 697 bp mRNA linear EST 01-AUG-2003
 DEFINITION CAB20006.11a_Fa_C06 Cabernet Sauvignon flower bloom - CAB2 Vitis
 vinifera cDNA clone CAB20006.11a_Fa_C06 5', mRNA sequence.
 CP210202
 ACCESSION CP210202.1 GI:33404575
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 697)
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and
 Cook, D.
 TITLE Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon',
 berries at various developmental stages
 JOURNAL Unpublished (2003)
 COMMENT Contact: Douglas Cook, PhD
 CABS Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 FEATURES
 source
 1..697
 /organism="Vitis vinifera"

/mol type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db xref="taxon:29760"
 /clone="CAB20006.11a_Fa_C06"
 /sex="Hermaprodite"
 /dev_stage="Bloom"
 /clone_11b="Cabernet Sauvignon flower bloom - CAB2"
 /note="Organ: Flower - Bloom; Vector: pDR; Site 1: SfiI;
 Site 2: SfiI; CAB2 is a cDNA library of Vitis vinifera cv.
 'Cabernet Sauvignon' clone 8 berries. Samples were
 collected at full bloom (80 to 100% flowers showing
 dehiscence of calyptras or caps and anthers fully
 extended). Sampled vines were located at the University of
 California, Davis, Experimental Vineyard. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAAGCAGAGTGGCCTTACGCGCGG-3' and
 5'-ATTCTAGAGCGCGAGCGCGGCGGCGG-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 22.8%; Score 128.8; DB 14; Length 697;
 Best Local Similarity 55.3%; Pred. No. 1.5e-25;
 Matches 270; Conservative 0; Mismatches 217; Indels 1; Gaps 1;

54 ACACACCCCTCTCTGTTTAAAGATGAAGCCGGAATGCTAACCCGCAATTCGCGACT 113
 |||
 45 ACACACTATCCATCAGTGTGTAATGATGCTCTGAGTCTTATCTTGTACTGGGGT 104
 |||
 114 ATTTCGCCCGGCTGTTTAAATGAGAGCTTGGCGGTGGGGTGGGGAACAGGGGA 173
 |||
 105 TTTTCCAGAGAGGATATTAACATCCAGCTGCGCATTTGTATGATGAGAGG 164
 |||
 174 CGTTTCCCGCATACCAATGCTGTGAGGCGGAGTGAACACCATGCAACATGACCAA 233
 |||
 165 TCTCTCGCATCAACTGTTGTTCTCGTGAACAGATGATCTTAAACAAAGTTGGTGA 224
 |||
 234 GGAATCTACAGTGTGTTAAGTAAATTAAGTACAGACATCCGAACTCCCTGTGT 293
 |||
 225 ACACACTTAAAGGCTAAATGATCTTCATGATGTTGAGATTTACCACTTACCTTTGC 284
 |||
 294 GGAAGGGAATTGATGCTGTGTAAGGAGGCGCAATGCCCTTACCGAGGGAATGAT 353
 |||
 285 TGAGCGAAGTGTGATGTTGATTAAGATGCTGTAATGCTGCTGGGTATGCTCT 344
 |||
 354 TGAGCTAGCCAGGATTTCCGGGCGCCGATTTGATGATCTCGGAACACCGTACCAT 413
 |||
 345 TGATATTTGCCAGCATTTTTCAGGGGCAAGCCATTTGATGATCACAATTAACCTT 404
 |||
 414 CGA-ATGCTGGGAGACCCGGGTAATGTAAGCAATCTCCAGATGTTGGCAAGTTGGC 472
 |||
 405 TGAGCTTACAGGAGCTTTAGCAAAATGTTGCTCTGCAAGGTTGTTAAGCCCTATGG 464
 |||
 473 ATTAAAGAGTGTCTGCAACGGGCAAAATTTGCTTGTGCGGGAATCCGGCTCAATAGC 532
 |||
 465 CATGTGAGGTGGCAGCAACTGGGCGGTGGCGTGTGCTGATGCTGCTGCTGCTG 524
 |||
 532 GGAATATCT 540
 |||
 525 AATTATCT 532
 |||

RESULT 13
 CD429148
 LOCUS 548 bp mRNA linear EST 03-JUN-2003
 DEFINITION ETH1_2_A06.G1_A002 Ethylene-treated seedlings Sorghum bicolor cDNA
 clone ETH1_2_A06_A002 5', mRNA sequence.
 CD429148
 ACCESSION CD429148.1 GI:31344791
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 548)
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Salzman, R., Chua, J., Gonzalez, W., Lane, S., Miller, W., Nanda, P., Olaseinde, O., Eastman, A. and Pratt, L.H.
An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid (ACC)-treated seedlings
Unpublished (2003)

TITLE

JOURNAL
COMMENT

Contact: Cordomier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University, sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: Sug5 (CTTTCGCTTAAGCTGCG).

FEATURES
source

Location/Qualifiers
1..548
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTX623"
/db_xref="taxon:4558"
/clone="ETH1_2_A06_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/note="Vector: pME18S-Fl3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from seedlings grown in hydroponic culture. At 8 days of age, medium was supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic acid (ACC) to induce endogenous ethylene (ETH) production. Roots and shoots were harvested after 27 and 72 hr and material from both time points was combined prior to RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-Fl3 vector (5-prime DraIII site is CAGCTGTCG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 22.8%; Score 128.6; DB 14; Length 548;
Best Local Similarity 57.7%; Pred. No. 1.5e-25;
Matches 267; Conservative 0; Mismatches 194; Indels 2; Gaps 2;
55 CACACCTCTCTGTTTAACTGAAGTGAAGCCGAGTCTAACCCGATTGCCGACTA 114
86 CATACCTCTCCATCTTGTGAATGACTGCTCTGTCCTCAACATTGTAAACAGAGTTC 145
115 TTGGCCCGCGGTGTTTAACTGAAGTGAAGCCGAGTCTGCCGGAACAGGGGAGC 174
146 TTTCCTCGAGGGGCTACAAATACAGACCTTCTGTTGGCCCACTGAGAGGAGGC 205
175 GTTTCGCCGATCAACCATGTTGTCGGGGATGAACAACCATGAAACAATGCAAG 234
206 ATTGCGCATTAACAACATGTTCTCTGTACTGTGAATCATTAAGAGAGTTAGTAG 265
235 CAACCTCAAGTTGGTTAAGTAAATTAAGTGAAGGAGCATACCGAAGTCCCGTGTG 294
266 CAGCTTAAAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
295 GAAAGGAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
326 GAAAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385

Db

473 ATTAAGAGTGGCTGGAACGCGCAAAATTTGCTTTGTCGCGG 515
506 ATCTGGAGGTGCGCAGAACTGAGACAGTGGACTGTCGCGG 548
RESULT 14
BG581436
LOCUS
DEFINITION
EST483170 GVN Medicago truncatula cDNA clone pGVN-64H22 5' end,
mRNA sequence.
ACCESSION
BG581436
VERSION
BG581436.1 GI:13596500
KEYWORDS
EST.
SOURCE
Medicago truncatula (barrel medic)
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE

AUTHORS

1 (bases 1 to 782)
Fedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gant, G.S., Town, C.D., Van Aken, S., Utecher, T., Cho, J., and Fraser, C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago truncatula, 2001

TITLE

JOURNAL

Unpublished (2001)
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M382496e TIGR sequence name: M382496e More information is available at: <http://www.medicago.org>
Seq primer: SKmod (CTA GAA CTA gta gat CC).

COMMENT

FEATURES
source

Location/Qualifiers
1..782
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-64H22"
/tissue type="N2-fixing root nodules"
/dev stage="effective root nodules harvested one month post inoculation with Sinorhizobium meliloti"
/lab host="E. coli strain XLOLR"
/clone lib="GVN"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from effective root nodules harvested one month post inoculation with Sinorhizobium meliloti. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using GigaPack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XLOLR cells."

ORIGIN

Query Match 22.8%; Score 128.6; DB 12; Length 782;
Best Local Similarity 59.1%; Pred. No. 1.1e-25;
Matches 277; Conservative 0; Mismatches 184; Indels 8; Gaps 3;
51 GAAACACACCTCTCTGTTTAACTGAAGTGAAGCCGAGTCTAACCCGATTGCCG 110

Db 228 GAGACACAGATTGGTGGTTCGTGATGATGAGATGGAATGATTACCGAATAGCTGG 287
OY 111 ACTATTTGCCCCCGCTGGTTTAACTATGAGAGCTTGGCGGTGGGGTGGCGAACAAGG 170
Db 288 AGTGTTCGGAAGAAAGAGATACACATTGAGTCATTTAGCTGTGGTTGAATCAAGTAG 347
OY 171 GGACGTTTCCCGCATCACCATGTGTGGTCCCGGGGATGAGAACCATCGAACACTGAC 230
Db 348 GCGCTTTT-----TCACCATGTGTGTTCTGTGACTGATAGGGTTTACGTCAAGTTGT 401
OY 221 CAAGCACTCTACAGATTGGTTAAGTAAATTAAGTACAGGACATCACCGAAATCCCTG 290
Db 402 TGAACAGCTTCAGAACTGTTCATGTTTGAAGTGAAGATCTTTCAGCTGACCTCA 461
OY 291 TGTGAAAAGGAATTGATGTGTGTGAGGTGAGCGCCCAATGCCCTTAACGAGCGAAGT 350
Db 462 AGTAAACGTGATGATGTGTGTGAGGTGAGCGCCCAATGCCCTTAACGAGCGAAGT 521
OY 351 GATTAGCTAGCCGAGATTCGCCGGCCCGCATTTGTGATATCTCCCAAGACACCGTAC 410
Db 522 TAACTGTGTGTGACACCTTCAGAGCCCAAGATGTGACATCTCGGACATTCCTGAC 581
OY 411 CATCGA-ATGATGGGGGACCCCGGTAATAATGTAGCAATCC-TCAGATGTTGCCAAGT 468
Db 582 GATTAGGTTACTGGAGACCCAGGGAAGATGTTGCACTCCAGAAATTTCCGCAAGT 641
OY 469 TGGCATTAAGAGGTGCTCGAACGGGCAAAATGCTTTGTTGTCGGGAA 517
Db 642 TGGAAATTAAGAAATAGCTAGAACCCGGAAGATTGCAATTAAAGAGGAA 690

RESULT 15
CK243942 891 bp mRNA linear EST 12-DEC-2003
DEFINITION EST127579 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POCa635 5' end, mRNA sequence.
ACCESSION CK243942
VERSION CK243942.1 GI:39789037
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 891)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: EST127578
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from TIGR via potato@igr.org
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
Source
1..891
Location/Qualifiers
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultiyar="Kennebec"
/db_xref="taxon:4113"
/clone="POCa635"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_id="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 22.7%; Score 128.2; DB 14; Length 891;
Best Local Similarity 57.8%; Pred. No. 2.5e-25;

Matches 269; Conservative 0; Mismatches 188; Indels 8; Gaps 2;
OY 55 CACACCTCTCTGTTTAAAGTGAAGCCGAGAGCTAACCCGATGGCCGAGCTA 114
Db 310 CACACCTCTCTGTTTAAAGTGAAGCCGAGAGCTAACCCGATGGCCGAGCTA 369
OY 115 TTTGCCCCGCTGGTTTAACTGAGAGCTTGGCGGTGGGGTGGCGGAAACAGGGGAG 174
Db 370 TTTGCTCGAAGGGGTTTAACTGAGAGCTTGGCGGTGGGGTGGCGGAAACAGGGGAG 429
OY 175 GTTTCGCCGATACCATGTGTGTGCGCGGGGATGGAACACCATGGAACACTGACCAAG 234
Db 430 CTTT-----TTACCATGTGTGTGTGCGCGGGGATGGAACACCATGGAACACTGACCAAG 483
OY 235 CAACCTTACAGTTGTTTAACTGAAATTAAGTACAGACATCACCGAAATCCCTGTGTG 294
Db 484 CAGCTCAACAGCTTGTGAATGTCTGGAAGGTGAGGATCTGTCAAGGAACACCAAGGT 543
OY 295 GAAAGGAATTAATGTGTGTGAGTGAAGCCGCAATGCCCTTAACCGAGCGGAAGTATT 354
Db 544 GAACGGAATTAATGTGTGTGAGTGAAGCCGCAATGCCCTTAACCGAGCGGAAGTATT 603
OY 355 GAGCTAGCCAGATATTCGCGGCCCGCATTTGTGATATCTCCGAAAGACCGTCAACATC 414
Db 604 TGGTTGTGGAATCTTTCAGAGGAATTTGTGATATATCAGAAACATTTCTGACTATT 663
OY 415 GA-ATGTTGGGGGACCCGGTAAATGTAGCAATCTCCAGATGTTGCCCAAGTTGCC 472
Db 664 GAGGTAACTGTGATTCAGGGGAAATGTGCTGCTTCTGAGAAATTTAAGCAAGTTGGG 723
OY 473 ATTAAGAGGTGCTCGAACCGGCAAAATGCTTTGGTGGCGGAA 517
Db 724 ATTAAGAGGTGCTCGAACCGGCAAAATGCTTTGGTGGCGGAA 768

Search completed: July 25, 2004, 07:28:13
Job time : 1616.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2004, 04:04:07 ; Search time 2268.95 Seconds

(without alignments)
10792.997 Million cell updates/sec

Title: US-09-893-033-17

Perfect score: 565

Sequence: 1 gfggaatttaccacaaatg.....ccctgaatccaagttrtag 565

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Database :

Listing first 45 summaries

GenEmbl:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_hum:*

17: em_hum:*

18: em_in:*

19: em_mtl:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pac:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_scs:*

28: em_un:*

29: em_vl:*

30: em_hg_hum:*

31: em_hg_inv:*

32: em_hg_other:*

33: em_hg_mus:*

34: em_hg_pln:*

35: em_hg_rod:*

36: em_hg_tam:*

37: em_hg_vrt:*

38: em_sy:*

39: em_hgo_hum:*

40: em_hgo_mus:*

41: em_hgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	543	96.1 105795	1 SYCCPNC	D64001 Synechocyst
2	272.2	48.2 342850	1 AP003597	AP003597 Nostoc sp
3	258.8	45.8 300450	1 AP005371	AP005371 Thermosyn
4	235.6	41.7 299350	1 AP006572	AP006572 Gloeobact
5	211	37.3 349652	1 BX569690	BX569690 Synechoco
6	207.8	36.8 7550	1 D63676	D63676 Cyanidium c
7	195	34.5 8200	8 D63675	D63675 Cyanidiesch
8	195	34.5 149987	8 AB002583	AB002583 Cyanidios
9	179.6	31.8 191028	8 BP038804	BP038804 Porphyra pu
10	177	31.3 6982	8 AF233069	AF233069 Galdieria
11	176.4	30.7 346683	1 BX572098	BX572098 Prochloro
12	173.2	30.7 164921	8 AF022186	AF022186 Cyanidium
13	161.8	28.6 302339	1 AE017164	AE017164 Prochloro
14	159.6	27.0 121524	8 AF041468	AF041468 Guillardia
15	152.4	27.0 5592	1 AF175526	AF175526 Streptomy
16	146.2	25.9 348986	1 BX572093	BX572093 Prochloro
17	141.8	25.1 4210	1 MSGILVB	L49392 Mycobactere
18	139.4	24.7 308050	1 SC039124	AL939124 Streptomy
19	138.4	24.5 15939	1 AE007128	AE007128 Mycobacte
20	138.4	24.5 24613	1 AY116644	AY116644 Streptomy
21	138.4	24.5 318050	1 BX248344	BX248344 Mycobacte
22	138.4	24.5 348676	15 BX842581	BX842581 Mycobacte
23	137.2	24.3 11759	1 AE011091	AE011091 Methanosa
24	133.8	23.7 11932	1 AE002508	AE002508 Neisseria
25	133.8	23.7 326301	1 NM622491	AL162757 Neisseria
26	133.8	23.7 349980	6 AX044033	AX044033 Sequence
27	133	23.5 300272	1 AE017213	AE017213 Geobacter
28	131.4	23.3 347137	1 BX640448	BX640448 Bordetell
29	131.4	23.3 349028	1 BX640433	BX640433 Bordetell
30	129.8	23.0 349305	1 AP005032	AP005032 Streptomy
31	129.6	22.9 281450	8 AK058418	AK058418 Oryza sat
32	129	22.8 1883	8 AE013291	AE013291 Methanosa
33	128.4	22.7 11318	1 AE015669	AE015669 Shewanell
34	126.6	22.4 10818	1 AY142799	AY142799 Helicobact
35	126.4	22.4 570	1 AE016912	AE016912 Chromobac
36	122.4	21.7 300774	1 AE001995	AE001995 Deinococc
37	121.4	21.5 10625	1 AL646068	AL646068 Ralstonia
38	120	21.2 13104	1 Continuation (11 o	Continuation (11 o
39	118.4	20.8 110000	2 MLF637	MLF637 Mycobacteri
40	117.4	20.7 44882	1 MLF637	MLF637 Mycobacteri
41	117.2	20.7 348950	1 BX321860	BX321860 Nitrobac
42	117.2	20.4 303050	1 AE014648	AE014648 Bifidobac
43	115	20.2 13257	6 AX492782	AX492782 Sequence
44	114	20.2 349980		
45	114	20.2 349980		

ALIGNMENTS

RESULT 1

SYCCPNC 105795 bp DNA linear BCT 04-JUL-2001

LOCUS Synechocystis sp. PCC 6803 DNA, complete genome, section:20/27,

DEFINITION D64001 AB001338 BA000022

ACCESSION D64001.1 GI:1001102

VERSION

KEYWORDS

SOURCE

ORGANISM

Synechocystis sp. PCC 6803

Synechocystis sp. PCC 6803

Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

REFERENCE

1 Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N., Sugiura, M. and Tabata, S.

Sequence analysis of the genome of the unicellular cyanobacterium

Pred. No. is the number of results predicted by chance to have a

	Query Match	Best Local Similarity	Score	DB 1:	Length
Db	36911	GTGGAATTTTAAACCCCAATGCGCACCGCGCGATCGCTCTTTGCCCCCATGAACACACC	96.1%; 99.6%; 0; 0; 2; 2;	Matches 565; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	60
Qy	1	GTGGAATTTTAAACCCCAATGCGCACCGCGCGATCGCTCTTTGCCCCCATGAACACACC			60
Qy	61	CTCTCTGTTTAACTTGAAGATGAAGCCGGAGTGTAAACCCGCAATTGCGCGACTAATTGGCC			120
Db	36851	CTCTCTGTTTAACTTGAAGATGAAGCCGGAGTGTAAACCCGCAATTGCGCGACTAATTGGCC			36792
Qy	121	CGCGGTGTTTAACTTGAAGATGAGCGCTTGGCGATGGGCGCGGAAACAGGGGAGACGTTTCC			180
Db	36791	CGCGGTGTTTAACTTGAAGATGAGCGCTTGGCGATGGGCGCGGAAACAGGGGAGACGTTTCC			36733
Qy	181	CGCATCAACCATGGTGTGTGCGGGGAGATGAGAACCATCGAACACTGACCAAGCAACTC			240
Db	36731	CGCATCAACCATGGTGTGTGCGGGGAGATGAGAACCATCGAACACTGACCAAGCAACTC			36672
Qy	241	TACAGTTGGTTAAACGTAATTAAGTACAGAACATCACCGAAACTCCCTGTGTGGAAGG			300
Db	36671	TACAGTTGGTTAAACGTAATTAAGTACAGAACATCACCGAAACTCCCTGTGTGGAAGG			36611
Qy	301	GAAATTGATGCTGTGAAGGTGAGCGCCCAATGCCCTTAACCGAAGCGAAGATTTGAGCTA			360
Db	36611	GAAATTGATGCTGTGAAGGTGAGCGCCCAATGCCCTTAACCGAAGCGAAGATTTGAGCTA			36551
Qy	361	GCCCAAGTATTCGGGGGCGCGCATTTGGAGATATCTCGAAGAACACCGTCAACCATCGAA-TG			419
Db	36551	GCCCAAGTATTCGGGGGCGCGCATTTGGAGATATCTCGAAGAACACCGTCAACCATCGAA-TG			36492

Oy	420	GTGGGGGACCCGGGTAAATGATGAGCAATCTCCACATGTGGGCCAAGTT-GGCATTAA	478
Db	36491	GTGGGGGACCCGGGTAAATGATGAGCAATCTCCACATGTGGGCCAAGTTCCGCAATTAA	36433
Oy	479	GAGGTGCTCGAACCGGCAAAATGCTTGTGTGTCGGGAAATCCGGCGTCAATACGAAAT	538
Db	36431	GAGGTGCTCGAACCGGCAAAATGCTTGTGTGTCGGGAAATCCGGCGTCAATACGAAAT	36372
Oy	539	CTGAATCCCTGGAAATCAAGTTTAG	565
Db	36371	CTGAATCCCTGGAAATCAAGTTTAG	36345
RESULT 2			
LOCUS	AP003597	342850 bp	DNA linear BCT 28-NOV-2001
DEFINITION	Nostoc sp. PCC 7120 DNA, complete genome, section 17/19.		
ACCESSION	AP003597	BA000019	
VERSION	AP003597.1	GI:17133729	
KEYWORDS			
ORGANISM	Nostoc sp. PCC 7120		
SOURCE	Nostoc sp. PCC 7120		
REFERENCE	1	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.	
AUTHORS	Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., Watarabe, A., Iizuguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Konara, S., Matsumoto, M., Matsuno, A., Muraki, A., Nakazaki, N., Shiino, S., Sugimoto, M., Takazawa, M., Yamada, M., Yasuda, M., and Tabata, S.		
TITLE	Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120		
JOURNAL	DNA Res. 8 (5), 205-213 (2001)		
MEDLINE	21595285		
REFERENCE	2	(bases 1 to 342850)	
AUTHORS	Kaneko, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research, Yana 1532-3, Kisarazu, Chiba 292-0812, Japan		
	E-mail:kaneko@kazusa.or.jp, URL:http://www.kazusa.or.jp/cyanobase/, Tel:81-438-52-3935(ex.2398), Fax:81-438-52-3934)		
FEATURES	Location/Qualifiers		
SOURCE	1. 342850		
	/organism="Nostoc sp. PCC 7120"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:103690"		
	/note="synonym:Anabaena sp. PCC7120"		
	/complement(174. 1478)		
	/gene="all4593"		
	/complement(174. 1478)		
	/gene="all4593"		
	/note="ORF ID:all4593 hypothetical protein"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="BAB76292.1"		
	/db_xref="GI:17133730"		
	/translation="MDKSTICLVDELPSDNTVKVNLADLYVAPGOMNLVGFNDTIR AVTGEIDAVIQRIDRATSLYNDPOQGYQSAVKYQRTDRDPTMATATLANKYGEK TGLSTLSNITTRADYDTSIDVSKAIFITFYCKLNGIPQNPQEPFRANSLANNQDA SLIRWALVCLGDIPLGSDPLSKIQGISGHSNITQNPFLVNSLPSNSPTDK LGIINGFNAVOGMMNLTVKGITQGISSHLGNFIQTADNDLQVAAFLOQTNYF EHGIGSVARLILKAHNLVKEIKQETIQPIQDSSSAVKSTQSAVKLEWDEE EBEWYATIEKVDQDFPINYIGSSHNEMWKGKDILATDYAASSDANGYAAGKVKC MSSEETWATETLTKIQGNQFYIHTGYDSSVDENWSDDIS"		
	1680. 1680		
gene	/gene="asr4594"		
	1680. 1680		
CDS	/gene="asr4594"		
	/note="ORF ID:asr4594"		

```

hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76293.1"
/db_xref="GI:17133733"
/translation="MTAIAPEKSPROVLLYLGERDNYFVEVPSLPGCLISQKTRBEA
LTNIEVLQDKSELTPEYQFINI"
1877..2770
/gene="a1r4595"
1877..2770
/note="ORF_ID:a1r4595
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76294.1"
/db_xref="GI:17133732"
/translation="MSKKLVDRLLPFTFPLGPTTTHNVQASAPFMEFNTVIDA
VQPNGDMVITQKYNFRDYNHLSRYIPLDNDRTTEVSVSGOLLSTISTSG
NHQILISWEHQLQSESHFVLTKEVIGGLVNNHYAQVWEALFSDRKAFLQAQVR
VEPEELTNKIKIFQSPGVADIRVDTKITEAVTKQPIEGTGLQVIFDRAGTKI
ETSGMSTQPFYERLWMLLGLLVFGIPLMLIVSRNGTSSSIDSDIDYGGGG
DYGGDYGGDYGGGGGGCD"
3082..4473
/gene="a1r4596"
3082..4473
/note="ORF_ID:a1r4596"
/transl_table=11
/codon_start=1
/product="malic enzyme"
/protein_id="BAB76295.1"
/db_xref="GI:17133733"
/translation="WADITPNSFSFVTLRIQINRVGMASITQALSSGCGTGLD
IEQTRQVSLRDLTVDASTEHAETTVQAVKAIPIKLVSYADRTFNHARGKISTSR
IPLKSVSDLAAYTPGVGRVCTAIAODPSQVYNTIIONTVAVITDGSAAVLGNGNFR
HAALFVMEGKAMLFKEPAGLDAPICLATOPEEIIQAVNKIAPFGVAMLEIAFR
CFEIEKRLDELIDIPVHDQGTAVITLALFNSLKVQSLADIRIVNGAGAGV
AIALIRLRAGAEEKIMCDSKGIISTRTDNEEKQEPAAKQGLAGAVGAVFTGV
SAPGILTPKWKSKAKDAIVRANPPEIPELISMDVAVIANGRPDYNQINNTVLA
FPGVFGALDCAQTITTTMTLEASALASLVNPSDLREHIITPSVDEVRVAVAAA
VQRAREGIARG"
5270..6853
/gene="a1r4597"
5270..6853
/note="ORF_ID:a1r4597"
/codon_start=1
/transl_table=11
/product="type I restriction-modification system DNA
methylase"
/protein_id="BAB76296.1"
/db_xref="GI:17133734"
/translation="VTEKAKQNNSSSEFEQCLWKADKAKNIDAAFYKIVIGLITFL
KTYSDAFELHOKLLAGBGDYAGNPEDRDYSKENVFVPEARMSYLDGQAKQDPI
GKTYDLAMEALBOENAKRLKILPRVYGOQLDQKSGGLDGLSITLIDAEKQAD
VLGRVYEYVTLQFALABGKGGQFPESIVKLLVEMLEPNGRVDPCCGSGMFPQ
SEKFKVNHQGRLDISIVGQSNETTVKLCMLNLAIGIDGSIKNPDESFLNDAK
DKLADFIANPNDSDWGGELLRNDGSRDLDXLPVPGNANFAMVSHFTYHAPTS
AGPVLISNGSLSSNMSGEDIKKALVQDLYDCTMLPTQLFYNTGIPACMLPISRYN
GKAKRDRGGEVLTIDASELGYMNRRSSAATREATISKIADPTHEMKSGGSYADICF
CKASIAIEKXNFVLTFGRTVGIIPDEVEDGVSELEKMSLTMALGEQMEGQULDEE
IKQQLKVGFIVSDELVDV"
6858..7121
/gene="a1r4598"
6858..7121
/note="ORF_ID:a1r4598
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76297.1"
/db_xref="GI:17133735"

```

```

/gene
CDS
/translation="MNMTOVILKLNIPVIEKLKHLAASHORTLEEBITSILEDVTE
TPITISKSRDWSPGFEQTCAGWQELLVREPOEADOREPL"
7226..7804
/gene="a1r4599"
7226..7804
/note="ORF_ID:a1r4599
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76298.1"
/db_xref="GI:17133736"
/translation="MIAVKKKPKLTPEEYFIWEKQOLRHEYLNGEYVANTGTONH
GRINSLIFIVKGLRGGCCQVNSDCGINIFERKDYVPPVYSTCDBRDTALQALO
YPCILVEVLSPTASVDRGDKFLYRNPSPQDYLLVDAEKIALIDLTKNRGNWEIF
NYQLGDNIKLOSIDISPIBSEYEDIVEESV"
7855..8145
/gene="a1r4600"
7855..8145
/note="ORF_ID:a1r4600
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76299.1"
/db_xref="GI:17133737"
/translation="MIFKRNMTQVILKOLNPIYKIKRLAASHORTLEEBITSILE
DVAKEEERKSEGEQWDTLQFERMQQENITFPDADPADIRDSVGRDVAF"
8211..8573
/gene="a1r4601"
8211..8573
/note="ORF_ID:a1r4601
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76300.1"
/db_xref="GI:17133738"
/translation="WVKLTQHOTETATGVVMEHLFGCLRLVESQKRRLLEYINQ
IPLKMTILNIDLKAKQMAQEBARLSKMGKXIPAFIDQIASIAVSNLILVTNNVSDF
BSFNDAVENWTVNSGEG"
8577..9749
/gene="a1r4602"
8577..9749
/note="ORF_ID:a1r4602"
/codon_start=1
/transl_table=11
/product="type I site-specific deoxyribonuclease chain S"
/protein_id="BAB76301.1"
/db_xref="GI:17133739"
/translation="MSEWKETTLGEIADIINGOSPTEGTCNNNGQLPLNGPTEFGD
RNPILPTQFTIDPKLIAAGADLFCVRSSTGRMNAWADQKVAIGIGIASIRADGILFQ
PYIRALITEKEKSLIAVANGSTPMPSIDHLLNLIYVLPSPNRIKISINLARIDDKI
YULRQONTLEIAIOTLFRHMTIDPEPNADGKPKYSGSGAMVSALGIVPEAMSVK
LQOYVINKHGVIFKGEYITTEYTEKILITLPNPKIKGGSFNDKSKYYISADDSNYVL
RKDAIATITWDSKRGDSLGPAPFIDIKGVLPFLNORIGVENNNDIKITLPIELCR
RYRSHIIGTSSGSTRHSPRICEVSFVLPDELIDKFSALATATIDKILISTL"
9944..10465
/gene="a1r4603"
9944..10465
/note="ORF_ID:a1r4603
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76302.1"
/db_xref="GI:17133740"
/translation="WIGKLRLENDERSFSHIALAESQAEKYNLISKYIDILFEV
KLESISILQYKNTLSNITDTRFLKFDGSHLKKRRHDLIVNNPFRITNIDNLNRELF

```

Query Match 48.2%; Score 272.2; DB 1; Length 342850;
Best Local Similarity 71.9%; Pred. No. 1,1e-65;

Matches	383;	Conservative	0;	Mismatches	148;	Indels	2;	Gaps	2;
QY	33	GGCTTCTTTGGCCCCGATGAACACACACCCCTCTCTGTTTATGTTAAGTGAAGCCGGAGT	92						
Db	39898	GGCTGTGTTGGCAAGATGAACTATCTTTCCGTTCTGTTAGTGAAGATGAGCCGGAGT	39957						
QY	93	GCTAACCCGATTTGCGGACTATTTGCCCCGCGTGTGTTTAACTTGAAGCTTGAGCTT	152						
Db	39958	TCTGTCCCGCATTTGCTAGTTTATTGCTGCTGCGCGCTTATATAGAAAGCTGCTGT	40017						
QY	153	GGGTCGCGGGAACAGGGGAGCGTTCCCGCATTCACATGCTGCTGCGCGGAGTGAAGA	212						
Db	40018	TGGTTCGGGGGAAGAGGGCGAGTTTCCGCAATTCATAGTGTGTTAAGCTGAGCATCG	40077						
QY	213	CACCATGGAACACTGACCAAGCACTCTACAAAGTTGTTAAGTAACTTAAAGTACAGGA	272						
Db	40078	CGTATGGAACAACCTACCAACCACTATACAAAGCTCGTTACGTTCCAAAGTACAGGA	40137						
QY	273	CATCACCGAACTCCCTGTGTGGAAGGAAATTTGCTGCTGGAAGCTGAGCGCCATGC	332						
Db	40138	TATTAATGAACCTCTTGTGTGGAAGGAAATTTGCTTGAAGTGAAGCCACGACG	40197						
QY	333	CCCTAACCGAGCGGAAGTATGAGTACCCGAGTATCCGGGCGCCGATGTGATAT	392						
Db	40198	CAGAACCGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	40257						
QY	393	CTCCGAAGACCGTCAACATGAA-TGGTGGGGGACCCGGGTAAATGTGACAACTCT	451						
Db	40258	GGCGGAAGATCTTTTAACTTTGAAGTGTGCGGCACTCGTGAAGTGTGCGCATGT	40317						
QY	452	CCAGATGTTGGCCAG-TTGGCATTAAGAGGTGCTGGAAGGCAAAATGCTTGTGT	510						
Db	40318	CCAGGCTGTAACAAGATTGATGATGAGAAATGCGCGACCGTAAATGCGCTGAC	40377						
QY	511	GGGGGAATCCGGCGTCAATACGAAATATCTGAAATCCCTGGAATCCAAATTTT	563						
Db	40378	TGGTGAATCTGGGTCAATGAGTACTTAAATCCTTGAAGGAAAGGTTT	40430						

gene	1588..2898
126..1655	/note="BP-1"
126..1655	/gene="t1r0594"
126..1655	/gene="t1r0594"
126..1655	/note="ORF_ID:t1r0594"
126..1655	/note="ORF_ID:t1r0594"
126..1655	competence protein Comh homolog"
126..1655	/codon_start=1
126..1655	/transl_table=1
126..1655	/protein_id="BAC08146.1"
126..1655	/db_xref="GI:22294316"
126..1655	/translation="MPKQITIAEHHPLAFAVDQIOELIVATGTHOVGVYGVN
126..1655	REKVRALIRNAGSFPMGQITLVNLTADLRKEGSPDLPISITLAAAGVATDILGD
126..1655	HLPLGELSDGTLQPAVGLALALAAQAGITGLVVPATATBALVGLKVGCHTL
126..1655	AEVAFIHDSSSSSPATPSPCAAASPOFTVDKQVQYQARALEAAAGSHMLI
126..1655	FVGPDSGKTMARLRLPTLLPTFALELVTKHSVAGILKERGQILDEPPRSPH
126..1655	SASGALVGGGSYRPERGELSLARGLVLTDELTEPRGVDVLELROPLEDQVSIAR
126..1655	OSVFPAPQFTLVASTNPPCGYVGDVPOCTGSPROQVMAKLSGLIDRLDLOYS
126..1655	SLKPEMTQPIGDSATYRQVILARSACQRPAREBNVHCAOMQSHHLEOMCSL
126..1655	DDASIHLEPRAIQGLSASATDRLIVARTIDLDLCETITSAHVAELQYRTIDRL
126..1655	Q"
126..1655	1588..2898
126..1655	/gene="t1r0595"
126..1655	1588..2898
126..1655	/gene="t1r0595"
126..1655	/note="ORF_ID:t1r0595"
126..1655	probable sugar transporter sugar binding protein"
126..1655	/codon_start=1
126..1655	/transl_table=1
126..1655	/protein_id="BAC08147.1"
126..1655	/db_xref="GI:22294317"
126..1655	/translation="MKQILPMPMKPNTAOLTVCNRRGFWALLGLIALVLIGCTTA
126..1655	TOSQPIELFMHGVNPPRRVYLOKLVDRFNARHQIRHQALVYQGPQQLKILAA
126..1655	VGNAPADLWVNTPTTGQVDELGALRPDDMWQASPRQVSPALIPTRVGHVWS
126..1655	IPPAVNNLGIYRPSLFAAGITLPTSTWELQVQTLAAQGITPLLPYLGGEBAV
126..1655	FSMLPFPWMSGSLGETPERAHIDTPAAMALEFMORLARGALIASAPRGYELDF
126..1655	IRGVAAQITGPTLIGLOQSGVDVGLPWPALKTATPAALGAGNLVFRSTPREEEA
126..1655	LAFNLVYLSEEPQTEWALGTGYLPINERVTNDRVAFVAGQPALRVFLDQMAAQR
126..1655	PNRGVARSQTLGRALESYTLGKSTPKAVTTAAEQRMQMRPR"
126..1655	complement (2895..3905)
126..1655	/gene="t110596"
126..1655	complement (2895..3905)
126..1655	/gene="t110596"
126..1655	/note="ORF_ID:t110596"
126..1655	probable heme d1 biosynthesis protein NtrU"
126..1655	/codon_start=1
126..1655	/transl_table=1
126..1655	/protein_id="BAC08148.1"
126..1655	/db_xref="GI:22294318"
126..1655	/translation="MGVHIQAVENVATYIISQRLQKRPPLTLMLEPLFRNLACSG
126..1655	CGKIQHLEVLARKYLTPBECFRAVEEGADIVAPGSEPLIPOIDIVGLAARAF
126..1655	VYLCTGNLILLENLHKRSPYRFSVSHLDGLRBMHKKCARGVPTAYKALIKRAA
126..1655	KGRVYNTITTFGVDPKKEIQEFPDYSLSGUDGCMSPGXAEMAPDQHEPLREBT
126..1655	RAURLELAPRRQKKNVNNPFLDFIVGERKDTCTWGPMSISVLDNQPCYIL
126..1655	NEGHYKRWELIENITDMDKXGHSNGKRCQVHCGEATAMDAENPANNVPAAS
126..1655	Lf"
126..1655	4441..6447
126..1655	/gene="trne"
126..1655	4441..6447
126..1655	/gene="trne"
126..1655	/note="ORF_ID:t1r0597"
126..1655	/codon_start=1
126..1655	/transl_table=1
126..1655	/product="ribonuclease B"
126..1655	/protein_id="BAC08149.1"
126..1655	/db_xref="GI:22294319"
126..1655	/translation="MPKQITIAEHHPLAFAVDQIOELIVATGTHOVGVYGVN
126..1655	VLPGIDAFVNLGDSRNGFTIHSVDGLPLKTKATITLLAPQKRVLYQVWKEPFG
126..1655	TKGRLTGNLSLPGRYLVMPYGRGVLSRITNEARHRLALGILVPRAGKGLVR
126..1655	TEAGVEEARVLELLELQKWEITIQQAASRPOLLDYQDEPIPIIDYRVNAAIRALAKPRD
126..1655	RIYVDPDGVKRVYKHLKNNVNPVGLIDYQDEPIPIIDYRVNAAIRALAKPRD
126..1655	LPSCGYITIOPTALITIVDNGSGFTSATSRETVLWNTCEATAEIRLQKQNLINIGV
126..1655	ITVFTIMDSRPDQITLIFHRSKRLPDKRKPDAQISLGIATVETKRCQGNINIFL

RESULT 3
AP005371/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 FEATURES
 source
 gene
 CDS
 gene
 CDS
 gene
 CDS

AP006572.1 GI:35211683
 Gloeobacter violaceus PCC 7421
 Gloeobacter violaceus PCC 7421
 Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
 1
 Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H.,
 Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y.,
 Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N.,
 Shiino, S., Takeuchi, C., Yamada, M. and Tabata, S.
 Complete genome structure of Gloeobacter violaceus PCC 7421, a
 cyanobacterium that lacks thylakoids (supplement)
 DNA Res. 10, 181-201 (2003)
 2
 Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H.,
 Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y.,
 Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N.,
 Shiino, S., Takeuchi, C., Yamada, M. and Tabata, S.
 Complete genome structure of Gloeobacter violaceus PCC 7421, a
 cyanobacterium that lacks thylakoids
 DNA Res. 10, 137-145 (2003)
 3
 (bases 1 to 299350)
 Kaneko, T.
 Direct Submission
 Submitted (15-AUG-2003) Takakazu Kaneko, Kazusa DNA Research
 Institute, The First Laboratory for Plant Gene Research, 2-6-7
 Kazusa-kanetaru, Kisarazu, Chiba 292-0818, Japan
 (E-mail: kaneko@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/,
 Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)
 Location/Qualifiers
 1..299350
 /organism="Gloeobacter violaceus PCC 7421"
 /mol_type="genomic DNA"
 /strain="PCC 7421"
 /db_xref="taxon:251221"
 complement(1)..1224)
 /gene="g11120"
 complement(1)..1224)
 /gene="g11120"
 /note="ORF ID:g11120
 unknown protein"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC89061.1"
 /db_xref="GI:35211684"
 /translation="MKTLVGNPKVGERFMRNEKEINGLIQLDERKAHVLIAQRRIGR
 TSLMOAATIKDRYLQIDIOSHVVDAIAELAAITRSGSIWEKTTGIFSVNLNN
 AARQDSIQLSEIKVTLRTGMTGSMOKGQDLPKLAESKRPVIFLDEPVLIVSR
 LKSGYQWTPRPBQGTDAFLSMIRANSIKHEKRIKVYVSGISGLEPILRQVGSATIN
 HLTPSISIGMEPTVYKSCVEALAEBSILTFPGDALTFFVDTLGCVPYHVLFPQQR
 TVCGEGLTHIPKTLVAVEREMILARGALHLHERLKLVGPELHPTALFLTEPR
 AAVCGESAAAIQVCELEYKQPVYARBIILILBHDGYMQQSNITRFVSLVQEMWR
 AQSILFLVPASARGK"
 1511..1966
 /gene="g11121"
 1511..1966
 /gene="g11121"
 /note="ORF_ID:g11121
 hypothetical protein"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC89062.1"
 /db_xref="GI:35211685"
 /translation="MAEYDASALERAIISARAGAROLAAPILEARVRYDGRSTIVYH
 TLNTYGVWIPVRSIQGLAGATISQQLREVTVSPSTRTGIRPALDAVFELDLKIGIRE
 TERMWQDRTAAKAGSARSAAKINAAARANGKGRKQPHVVAPE"
 2128..3642
 /gene="g11122"
 2128..3642
 /gene="g11122"
 /note="ORF_ID:g11122
 probable long chain fatty acid CoA ligase"
 /codon_start=1

gene /transl_table=1
/protein_id="BACg9063.1"
/db_xref="GI:35211686"
/translation="MNLPVRAVYHGPRITAVTDQCFYQDILLSSQVARTILDG
ADLRSEQVNSLVAPGEVYATQWIGIRAGGIAPVLCVSHRPELEVITNSGASIVT
AHPSESLRPIIAEAHQRLFTLSLSLPPDPAQLPEVEPTRRALIYTSGETPKKCV
VTGNIICAOVTSILSMEWTSGDRIATHVPLIHIGIVVLTALWAGCEHMLSEF
DAEYVETMIGSNFTLIMFAPVTITKLIATWEXAGROMDSRGCAGMDLVGSA
LPVYLEWHTMISHFILPERYGMTEIMALSINPLHGPRAGVGMALPOVDRLVDS
GVSVPAGTPGSIQVNGCPVPLEYQKEXEATKAFQDSFCTGDIATVEQDCYRLIGT
SVDIITKTERGVASALEIEEVLRTHPDIOECAVGVADPEWGEVCGALVQSGGNL
EPRFWRAGYLAAYKVPFRLISVELPRNMGKTKVPAQDLFRAQOIASQPA"
complement(375. .5196)
/gene="g111123"
complement(3775. .5196)
/note="ORF_ID:g111123
unknown protein"
/codon_start=1
/transl_table=1
/protein_id="BACg9064.1"
/db_xref="GI:35211687"
/translation="MPSIATIGGSLAGQAAYOUKUGLVADYEAKSFVNGRIOSRK
NVNGEIVELIGFTFINSNRDITLQSENLPLFNPAKDQCPPEPETGYIPDGQNR
LESVEAEKLRPAQADALADLENDFKRAPQCDRLSVAEYLTHADKIPERITRL
IENSIRTYGVEPVDSALQLEFNLPVVDGAVFLVSSPEI FVEGSGESLILASLAA
VLPGIIRHMLATRLKKGCKYKRLTFAGAQRVEDPVI VAVPVAGLKKIIFOVLP
FRFRTKEVSLFENEKLPFAGSEVRVOREGFSWMTDGFSGVMEETOROPASDGA
LTFEGAGDEVNATNISGACTGEAFIORFEAVVDGAAATGRRSPRMNDPPIGGG
YTFPRGQTLTPGGFTYTESVITPAFRQDVNGLIPAGEHLSDEFYGFNGAOTGRL
AAQVRLQIQEGQLASGKPTSA"
5942. .6382
/gene="g111124"
5942. .6382
/gene="g111124"
/note="ORF_ID:g11124
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BACg9065.1"
/db_xref="GI:35211688"
/translation="MLSMDPVATRIIVDTSVFISALIGPAGSRNLRCLQGYOPL
MKNALFCYBESIVGRPHILAOCPISQNDILTMALMSVSMISITLYLMRNKLKED
NHLIELAVAGNAQVATINVRDPQVABLLFPDSLTKPEELIRR"
6389. .6607
/gene="gsr1125"
6389. .6607
/gene="gsr1125"
/note="ORF_ID:gsr1125
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BACg9066.1"
/db_xref="GI:35211689"
/translation="WATITRLPDDKHEIRLELAQTGDISVKNLIIELTSTIALAEFDA
YTRFAMAIGNPESGLTKLRDKLDALIG"
complement(6672. .7328)
/gene="g111126"
complement(6672. .7328)
/gene="g111126"
/note="ORF_ID:g11126
hypothetical protein"
/codon_start=1
/transl_table=1
/protein_id="BACg9067.1"
/db_xref="GI:35211690"
/translation="MRKQTRAMQPTPEVQDAMMSAESFPLATJDNPALVHAKYVYHL
GMRLKEMSPVGAHAEEDNNLLAQVVGIPAFSRGIRLKGVINLTARKTELQACQPDILAH
YENQSPPARINSPVDEJOIAPALAEIATLTGDTGKRELVQLOLGEVWVVDV
AENQVILFGPIPGVDSLEPIVASRVLPGTLAATLCEATLRIGSGEGDAAAIRVYLDLE
"
complement(7361. .8260)
/gene="chrb"

CDS	complement(7361..8260)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
-----	------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Query Match 41.7%; Score 235.6; DB 1; Length 299350;
Best Local Similarity 67.7%; Pred. No. 2.6e-55;
Matches 359; Conservative 0; Mismatches 169; Indels 2; Gaps 2;

QY	23	ACCGCGCATGCCCTCTTCTTGGCCCCCATGAAACACACCCCTCTCTGTTTGTGAAGATG 82
DB	15531	ACTATGGCGCATTTGAGAGCGAGCGTGAAGCATACCTTTCGGTATTGATGAAAGATG 15472
QY	83	AAGCCGAGTGTCTAACCCGATTCGCGACATTTGGCCCGCGTGTGTTTAACATTGAGA 142
DB	15471	AGGCGCGGCGCCCTACCGCGATGCGCGCANTTTCGCCCGCGGCGTTCAATATGAGA 15412
QY	143	GCTTGGCGGTGGGGTGGCGGAGACAGGGGAGCGTTTCCGCAATCAGATGCTGCTCCG 202
DB	15411	GCTTGGCGGTGGGGTTCAGCCGAGCAATCGGGCACACCCCCATCAGATGCTGCTGCGG 15352
QY	203	GGATGAGAACACCATGAGAACACTGACCAAGCAACTCTCAAGTGGTGAAGCTAATTA 262
DB	15351	GGGACGAGACAGCGCGTGAAGCAATTCACAAAGCTGCTGCAAGCTCATCAAGCTGCTG 15292
QY	263	AAGTACGAGCAATCAGGAACTCCTGCTGTGAAAGGAATTGATGCTGTGAAGTGA 322

DB	15291	GTGTGAGAGCAATCAACCAATTCCTCGTGTGAGCGGAACTGATGTCATCAAGATGA 15232
QY	323	GGCCCAATGCCCTTACACGAGCGGAGATGATGAGCTGACCGAGGATATTCGGGCGCCA 382
DB	15231	AGCCACACCGCCCGACCGCGACGAAATCATATGATGACCGCTCTTTCGGGCGCCA 15172
QY	383	TGTGATATCTTCGGAAGACACCGTACCATGGA-ATGCTGGGAGACCGGCTAAATG 441
DB	15171	TGTCGATCTGGCGGAAGATGTCACCGCTGAGATGACCGGCATCCGGCAAGATG 15112
QY	442	TGCAATCTCGAGATGTTGGCCAAAGT-GGCATTAAGAGAGGTGCTGAAACGGCAAA 500
DB	15111	TGCGCATATTCAGATGCTCAACCGCTTCGGCATTAAGAAATTCGCCGACCGCAAA 15052
QY	501	TTCCTTGGTGGGGAATTCGGGCTCAATACGGAATATCTGAATCCCTG 550
DB	15051	TTCCCTCACCGCGAGTGGGGGTGAACACCGAATTTCTCGCACCCAG 15002

RESULT 5
BX569690/c 349652 bp DNA circular BCT 23-SEP-2003
LOCUS
DEFINITION
Accession
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL

1. 349652
/organism="Synecchococcus sp. WH 8102"
/mol_type="genomic DNA"
/strain="WH8102"
/db_xref="taxon:84588"
complement(173..1291)
/locus_tag="SYNM0355"
complement(173..1291)
/locus_tag="SYNM0355"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="hypothetical"
/protein_id="CA806870.1"
/db_xref="GI:33638575"
/translation="MITTALPRMGSGRISEAKLQFADNHEVVAAMRQYVARIGPAP
VARSWLYALBESGIIITKGDFTWGSKWLADRKDELIFELVGLADSTALSGSDYDKE
ETPREYINRQLRBSLIDOKHYWPAWYKHQYEPFIWTERKDLKLFEPFLPNARIRF
ASKGQADVNSRVALIOECDMALGGLGEPVLYGDLDPMGVKSMDALAEMLRPLAKYL
NMDSGLDMDDGRIVRFGNAEFDVDAAGLVWIDGLFTSSGGLDANPNHHRKRPVYO
DYLETYGRKREANALANPRPARSLRDELIMOWLSHGKQWSENRKASSEATGIV
DAITRMALMFDVSGALYNPRLEQAVNSGLNSLPFAAN"
complement(1368..3358)
/locus_tag="SYNM0356"

2 (bases 1 to 349652)
Larimer, F. and Palenik, B.
Synecchococcus genome consortium
Direct Submission
Submitted (01-JUL-2003) Larimer, F., Submitted on behalf of the
Synecchococcus genome consortium, the DOE Joint Genome Institute,
Production Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA
94598, USA, and the Genome Analysis Group, Oak Ridge National
Laboratory, 1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov

Location/Qualifiers
1. 349652
/organism="Synecchococcus sp. WH 8102"
/mol_type="genomic DNA"
/strain="WH8102"
/db_xref="taxon:84588"
complement(173..1291)
/locus_tag="SYNM0355"
complement(173..1291)
/locus_tag="SYNM0355"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="hypothetical"
/protein_id="CA806870.1"
/db_xref="GI:33638575"
/translation="MITTALPRMGSGRISEAKLQFADNHEVVAAMRQYVARIGPAP
VARSWLYALBESGIIITKGDFTWGSKWLADRKDELIFELVGLADSTALSGSDYDKE
ETPREYINRQLRBSLIDOKHYWPAWYKHQYEPFIWTERKDLKLFEPFLPNARIRF
ASKGQADVNSRVALIOECDMALGGLGEPVLYGDLDPMGVKSMDALAEMLRPLAKYL
NMDSGLDMDDGRIVRFGNAEFDVDAAGLVWIDGLFTSSGGLDANPNHHRKRPVYO
DYLETYGRKREANALANPRPARSLRDELIMOWLSHGKQWSENRKASSEATGIV
DAITRMALMFDVSGALYNPRLEQAVNSGLNSLPFAAN"
complement(1368..3358)
/locus_tag="SYNM0356"

CDS
complement(1388..3358)
/locus_tag="SYNM0356"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAE06871.1"
/db_xref="GI:33638576"
/translation="MKPVSNWROKLLPOLTGPLIPGAGGKFKAPIDPSNGAKCH
QMSFTPBBIAMGKRVLCVGTRELSPDAGIYVCPIDGLSAISELIRKGIHQVTVW
VGRNIDSNRYKLFRRVPERKMEHLPGKEVADGKVEIFWSTGQCI VAGEHSQSG
EYVWQNRNPERIAPIDPMWALMOKAMAKTENRRSHKQVGMHDAIPCIIGRTPEPC
ROSADGKYLICHYGKRWSPVPWAKGETVYKGVWYACGDETAAGNAVRIIVEQQ
DLROKVOAGDALKLMAEQLGDPRLNLSRGIHNGAETATOTENLYLRLSMPSP
NRMSQIKARDAFIELAHFHOPDPVYVLSNLSADPLPEDMNHIGRFLFNVDLIDA
FMPRYLTSAVARVYKPGGCOGOTPCIIIGCGIGKTEAGRSI.RGHRFYCGDLSALDID
DYTRIQFVGMELGELNGITRTRTOQEKAKHLSRVDIVRKRYAPGTEPIKRSVFPA
TTNKSPLTNGSTFRVMI PVGDEKLPVQRLAEADALWTALREPFQNGLTWSTDE
MDRIILTRNSDYDLVDPMWSDLI.GDWLKRQTAAYVORQIYDFLEIGPERONNFNAORI
REIMTHGMLYDORRIGTGKWRFAFWNSRPD"
complement(1375..3638)
/locus_tag="SYNM0357"
complement(1375..3638)
/locus_tag="SYNM0357"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE06872.1"
/db_xref="GI:33638577"
/translation="WRALMEETASTTAATVAQAOLSPETIIMRDEPNPQGHDMQ
FGLGTRVYRWKKEQALQSLMWNKRESSAEVETGRTAPLTV"
complement(1641..3901)
/locus_tag="SYNM0358"
complement(1641..3901)
/locus_tag="SYNM0358"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE06873.1"
/db_xref="GI:33638578"
/translation="WQGPDRPIPIRPGRLTTHPEGHEHOLERLVRYTTRARVLDDEG
PSIGTLDGKNPQGRBKKANFMALQAEILHHEWNGTPT"
4202..4372
/gene="smta"
/locus_tag="SYNM0359"
4202..4372
/gene="smta"
/locus_tag="SYNM0359"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="Bacterial metallochionein"
/protein_id="CAE06874.1"
/db_xref="GI:33638579"
/translation="YSTAIKCACPKCTCMAEBESAIYLGKFPGSTSGSTHANNEFC
HGSGCGCKCGR"
complement(4431..5711)
/locus_tag="SYNM0360"
complement(4431..5711)
/locus_tag="SYNM0360"
/note="In addition to BLAST hits contains the applicable
substitutions in the R-H-R-Y motif identified by Nunes-Duby
et al 1998"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="weak similarity to phage integrase family"
/protein_id="CAE06875.1"
/db_xref="GI:33638580"
/translation="NAPDSDRPMRQRLVETTERMNGSISRYVDRIYLRKRGKRA
ETAYIPRENIMDEVHRNVAIKRIEAVYRAMNGEKTLAAVAADADHDSLDGGEKAKT

gene
6342..6740
/locus_tag="SYNM0361"
6342..6740
/locus_tag="SYNM0361"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE06876.1"
/db_xref="GI:33638581"
/translation="WIDLDAQIDQSYREASDAEVARRLERIEQIPGAQGLIPRKY
GTVNRKALQENLTLASLITQADPAALANYCGDASVKSMDEREAKRLVEALRMRT
ECLEARENSAAKTRBQOLVSGINPMTGRYF"
6740..7087
/locus_tag="SYNM0362"
6740..7087
/locus_tag="SYNM0362"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE06877.1"
/db_xref="GI:33638582"
/translation="WLAIRPAPCCGMTHQGPQDSRRLDRLQULTRDHYLZADSFSGMP
ASFVEETWDMLEKAVAPPRQLTAHVAELERQISLRETELQSGGLDQRDAAV
DLRQRLHLETS"
7667..8926
/locus_tag="SYNM0363"
7667..8926
/locus_tag="SYNM0363"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAE06878.1"
/db_xref="GI:33638583"
/translation="NALTLMBDLSGKFIHTPDLRKIPNAKRPBPCVEYELFASNS
KNAKGIISFIDGSLDMVDPDSALDESNLYGSPRYGSDTASNGSIQWLGKGDIGEN
AAGGSAVIQADVDWALSSREGSLANI.SDLYDLKNSLDDKFDTLTMIIGGIGNVINI
PDGENPVDWNSQAADQMESLINSILANDMSKIVIGYVISAHPNFIQILHNKFSY
LAENNSKVEFINTDYL.YDQAGADAVLRSL.FGQDNGHGVBEGRRLVYSPFLQGLQ
LKPAPPEPEYDSII.EBSVLGKGLKGTGTADAF.FPDSFEAFTKSADRIKIPNASGDT
IAYGNAPSPSIGVSDISFRASRSKKEPKOMKEDYDFYFEEKGLYFDGNAEAKNW
GNSSBEGVAILKGRPELTAEDTLIA"
8950..9330
/locus_tag="SYNM0364"
8950..9330
/locus_tag="SYNM0364"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE06879.1"
/db_xref="GI:33638584"
/translation="NAITTRAQOCRRNEALQLISSGVPPTDAASQLTVKMGCSRTS
LRDIEIAOSELIANALDVELQQVGVWLATQYORLAKAKARDQGYSAAGALNALRMY
VQDOLDAOSRAHRRGRFTHSYR"
9332..9691
/locus_tag="SYNM0365"
9332..9691
/locus_tag="SYNM0365"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE06880.1"
/db_xref="GI:33638585"

```

/organism="Cyanidium caldarium"
/mol_type="genomic DNA"
/strain="RK-1"

```

```
12. .83  
/gene="trnk"  
12. .83  
/gene="trnk"
```

```

/product="tRNA-Lys"
108. .836
/gene="tlypA"
108. .836

```

```

/gene="trpA"
/codon_start=1
/product="tryptophan synthase alpha subunit"
/locus_tag="b1400000000_1"

```

```

/db_xref="GI:2541888"
/translation="MFIAYLTGAPDLNTTKCALINLANDGAVIEIGVPSDPLADGOC
MLOKASQALKNRRLLEQTMHLAEIELEPVVILAYNOIFHYGEKWTTLVNLK
IKAT TUDRI QVERA AT IBAVMDUHHU HMTET ICDMDVTRAVKO AI AQDDWIVT VSDPT

```

```
GATGATDAHPDNOI PNMIAELKQATTTP IALGFGHQKQDLQVKS WGDGVII GTACN
QILLEKGVDTLEWISAMKSSYP"
800. .2023
/motA="OPFA07"
```

```

/codon_start=1
/protein_id="BAA22824.1"
/db_xref="GI:254189"
/translation="MKDKDKITIDVCAITAFI SKRQNIIDTEETIRVYETOGSMITIN

```

DILDELMKFOAKYGFEMKONNSSEEDVDIAINAVEERLDGEAVSNQVMS
 SITDRAFVLQMPKLAALVHLTQOIVDQIHHPKXASMVVESKRDACLGVVTHLLPR
 SILKIGDHIKAEAGSGKRIALMNVKQOQVYBAKAGELVHLTGFNOVLCLATAINKKAK
 ERENTSNQITHHGSKNLTYPSEQSGIENTVCEIYVOEKEMLSAEVGSITMDKE

OGT1AGCLVTKGVFQOTGMVCOJLLGNHLLKATAHINEIKIKHQEYKEVKASQJECGISMKH
LAOBTNSILVAMOMTARHKQVKEADFEOLITRLQOIQAQPISTIITSKAVKQVE
EWMNDQVUDEITLVN" 3095)
CComp]ement(2023

```

/gene="trnT"
complement (2023. . 2095)
/gene="trnT"
/product="tRNA-Tyr"

```

```
2096. .2674
/gene="Ips4"
2096. .2674
/gene="Ips4"
```

```

/codon_start=1
/product="ribosomal protein S4"
/protein_id="BAA22825.1"
/db_xref="GI:2541890"

```

STI.TV.TOTOPVOKEWJIA.FYVNNEJL.VVEEYTBRYGA"
/FIANSIATION="MSKRLDGPRAVYTRKQJALPALINKS.FNKKI.IAFGEHNNI.KRUS
EFAVOLQOKQLOYYIGITNSQJLAFPRQABRSKSTGLALITMLETRLDHLVYRAGF
APT.PPARQJLVNHGHYKNGKQJLTIASMACEVNHVIEVKSSSPPKPEYLPPLVQLS

```
complement (26/1. .3432)
/note="ORF253"
/codon_start=1
/protein_id="BAA02826.1"
```

```
/ad_xiel=gl:2541891
/translacion="MDDYIFREYRLSEPQPTWQSLMOWYVSPVOIQHLLLSRQVQSI
HLPIFKSKSCGVLDYQFHGYQFKSNHPKCHDAFROSLACPIYVVVNPKEQLILSPRKH
WSEYQAPTEMWVWLDCEVROSPSTSYLIFEHPDDAOMYLQSVI NNPERDKNSFVOH
```

```

GIDLFIFLKKFNKADKAVLVLFPHLQWQWQFUGLFFVFRVQVQKREWLFUSNQWQF
KRMKLSCKDCVVSLESILEADANFQARCVLVPLS"
complement(3608,.3913)
/gene="trxm"

```

```
complement(3508..3913)
/gene="trxm"
/codon_start=1
/product="thioredoxin m"
```

protein id="BAA2827.1" /db_xref="GI:2541892" /translation="MQIDELSPDTEVLSHQVLVDVFWAPWCGPCRMIAPIILDEIAQBEIDIKKVKNLTNDENPNLATEYGRISPTIMLEKGGORLDIVIGAVPKSTLNTNTEKLT"

4046. 351


```

/codon_start=1
/product="tryptophan synthase alpha subunit"
/protein_id="BAA22814.1"
/db_xref="GI:2541878"
/translation="MLIAYLTAGAPDINTTKEAVMKLAKKADVIBIGVPSYDPLAD
AIIQKASQALMNGFHLDELMLISEVIEIPLVILAYNOIMHYGEVWKVLAH
YVKGLIVDLPEYRSKTLROI CDRYGNIIIMLSPTNKTBAQLACADMIVYVSR
TGVTGLETPFPKQIPKILGELKRYTKRPIALGRISSISQVXKVMGADGVIISAC
MQLLEKGVDDLSWISVMKSFCEP"
936..2198
/feature="ORF420"
/codon_start=1
/protein_id="BAA22815.1"
/db_xref="GI:2541879"
/translation="MDKCDKILLISGTVADLAKOTKKEVTEPIRYAFNQSIRLN
DIVDELMLQCKTGWEMQJINIEDEFADEVADYARDEDEMEIAYSAYKQLQAIW
SNHWSIAIKGVFVLQMPKILKALANLTOQITDKVNHPRGAPLIVESKDKNLGIT
TTLIRKSKLVGVDVIBGNKIRAMQINORIKVAKGEIVHSGENVYSITNLTLY
SAKKGEKASNPQTNQKLCHEPKGNLIVRSSESSCYFLLEKQQLSAEVS
IRMODVLAQTSKILVAMQTCVNTNTVNPVPIKIEANFEDILQLESVAKDR
IVSKALVKQVFRGSGIAGCVLQGEFRTGEVIMRKQYQARARLSEIKIKQAVK
EVAPQEGIFMKDWTOMVDDELIVLN"
complement(2197..2270)
/gene="trnI"
/complement(2197..2270)
/gene="trnT"
/product="tRNA-Thr"
/feature="anticodon:UGU"
2271..2846
/gene="trpS4"
2271..2846
/gene="trpS4"
/codon_start=1
/product="ribosomal protein S4"
/protein_id="BAA22816.1"
/db_xref="GI:2541880"
/translation="MSRYLGRPVRIIRRLGILDPATFNKSPKRTGVPGHAKTRKL
SEVAGQLQKQKLYGGITNNQIARFPRQKSKRSASTGIELNKEIRLDHYVRG
FAPILPARQLVNHGVKNGNOVITIASFAQVNHIIIVKAKSPSSAQLPPYLQVENQ
FVKMIQPVKMDMLAFRVNELIVVEYTRVGA"
complement(2849..3595)
/feature="ORF248"
/codon_start=1
/protein_id="BAA22817.1"
/db_xref="GI:2541881"
/translation="MHDIYIERYRLTEDEPRKQWLLQWFSVPELEHLLLSQIQSV
CIPSKPESFIDYQESYSPKNTQLVHQNILASISLACPIYVVPKQELILSPQWE
LEKAPETLNVFEDHIFRNPTSYLAFHFDADAMYLKSVLVNPPYERDKNSFVEQD
LFSFLKLNRTDLKVLVPLRLQCKWSFEGTPIVFRVQWQDHEWLFFSKDAQQFIKN
QISCKITVDYISBSALNRQARCLVPLST"
complement(3801..4109)
/gene="trnM"
/complement(3801..4109)
/gene="trnM"
/codon_start=1
/product="chlorodoxin m"
/protein_id="BAA22818.1"
/db_xref="GI:2541882"
/translation="MLHIDELTFENEVIOSEKLYLVDFWAPWCGPCWIGIILEIK
BENLKVQVNTDENPMLATFYGISIPILMLFKKGQVDTVIGAVPSIILHTINKYL
4542..6008
/gene="trnL"
4542..6008
/gene="trnL"
/codon_start=1
/product="ribulosebiphosphate carboxylase large subunit"
/protein_id="BAA22819.1"
/db_xref="GI:2541883"
/translation="MAQSVQERTLKNRYSGVIPYAKQGYWNPVYVKSQDMLAF
RYTPQGVDPYIEAAAVAGSSTATVVTDLTACDVYAKAVYVDVPTAPDOVF
VYIATIDLFEBGSIANTLSTIGNVFEPKXKALRIEDNRIPIPSYKLTGQPGTGYI
VERERLNKFGRLGCTVFKPLGLSGKNGVAVYBGLGGDLFDLDDENINSQPFKSW

```

```

RDRYLVEGVNRAAASGEVKSGLVNLVNTATMEECYKRAEFAXEVSIIIMDLVIG
YTAIOTMAIWARNNMILHRAAGNSTYSRQKNGINERVISKWRRAVGHVHAGVY
VKELEGDPIIIGKPYNTLLPKLEVNLPGLFFEMDASLIRKTLPVASGGIHQOMQ
LHVLGDEVVLQFGGGTIGHPDGIQAGATARVALAMVLAARNRGROYFNNGPOLND
AAKNGPLOTALDLMKDAIFNYISTDTSDFETPLAN"
6027..6443
/gene="trnCS"
6027..6443
/gene="trnCS"
/codon_start=1
/product="ribulosebiphosphate carboxylase small subunit"
/protein_id="BAA22820.1"
/db_xref="GI:2541884"
/translation="MRVQGTFCVLPDLTDQIRKQIQYALNQMAVSVEYTDPHPR
NSYMMGLPLPEVDVDPATIMEVQECRQEPNHYIKVLAFNSTKGIESTALSPVNR
PANERGFYLRILIESNDROYRQYHISYAEKREGSRY"
6522..7409
/gene="trnQ"
6522..7409
/gene="trnQ"
/codon_start=1
/product="trnQ protein"
/protein_id="BAA22821.1"
/db_xref="GI:2541885"
/translation="MKMSVQTSQCTVNLCTEFKQTOIOWVLDLIDRELIGQVKTTRI
REIALLIVDRKQLSLSSNPGILHNSFTGSPGKTTVARKKADILYRGIKKGH
LITVTRDVLQVQYIGHTAPKTKQVILKNAAGVLPIDEAVYLRPNDEKDVGAELIIL
LOVMENQRDLVLIIPAGYKDMRFRYSNPGILARVANVAVPDPYEGMLMIGKIML
OEOOYQMTPEAEKVFLOYIQRMEQPHFANARSVNALDRARLQANRIFATPPKCLT
KFDLVTIOAEDILKSRFLQ"
7418..7488
/gene="trnG"
7418..7488
/gene="trnG"
/product="tRNA-Cys"
7494..7579
/gene="trnL"
7494..7579
/gene="trnL"
/product="tRNA-Leu"
complement(7586..8086)
/gene="trnVH"
complement(7586..8086)
/gene="trnVH"
/codon_start=1
/product="acetolactate synthase"
/protein_id="BAA22822.1"
/db_xref="GI:2541886"
/translation="MKHTLSVYVEDDEAGVLRIRAGLFARRSFNIESLAVGAQAGIS
RITNVVPGDRTIEOLMKQLYLPIILGVENLTVPCVERELIMKVOANSOTREIL
DMANTFRARVVIDSADVMLVETGDPKGVNLEQILTFGIVEVARTKIALKSSKY
NTENLR"

```

ORIGIN

Query Match 34.5%; Score 195; DB 8; Length 8200;
 Best Local Similarity 65.0%; Pred.No. 6e-44;
 Matches 319; Conservative 0; Mismatches 170; Indels 2; Gaps 2;

QY	DB	CGACTATTGGCCGCGCGTGTGTTTACATTGAGACCTTGGCGTGGGTGGCGGAAACA	107
QY	8087	CATGAAACACACCCCTCTCTGTTTGAAGATGAGCCGAGTCTAACCCGATTCG	107
QY	108	CGACTATTGGCCGCGCGTGTGTTTACATTGAGACCTTGGCGTGGGTGGCGGAAACA	167
QY	8027	GCGCTATTGGCCGCGAGTAGCTTAAATAGAAAGTTAGACGAGACCTGCAGAGACA	7968
QY	168	GGGGGACGTTTCCCGCATCAACATGCTGCTGCGGGGAGATGAGAACCATGGAACACT	227
QY	7967	AGCAGGTATATCTAGAAATTAACCATGCTGCTGCTGCGAGAGCGATTAAGCAATG	7908
QY	228	GACCAAGACACTCAAGATTGTTAACTAATTAAGTACAGACATCACCGAAACTCC	287
QY	7907	AATGAAGCAATTATATAAAGTATACCCATTTTACAGAGTGAAGAAATTTGACACAAGTACC	7848

```
QY      288 CTGNGTGAAGAGGAATTGATGCTGGTGAAGAGTGAAGCCGATGCCCTTAACCGAGCGGA 347
Db      7847 TTCTGTGGAAGAGGAATTGATGATGAAGGTGCAAGCCATCCCGACGACGAGAGAA 7788
QY      348 AGTGAATTGAGCTAGCCAGGTATTCGGGCGCCGATTTGGATATCTCCAGAACACCGT 407
Db      7787 AATTATTGATATGCTATATTTTCGGCGACGCTGTGGATATATCAGACAGACATGT 7728
QY      408 CACCATCGAA-TGGTGGGGGACCCGGGTAAATGTAGCAATCTCCAGATGTGGC-CA 465
Db      7727 CATGTTAGAAAGTAAACGGGTATCTGGAATAATGTAGCACTGAACAATTTTAACGAA 7668
QY      466 AGTTGCATTAAAGAGGTGGCTCGAACGGGCAAAATTTGCTTTGGTGGGAAATCCGGCGT 525
Db      7667 ATTGGATATGTGAAGTACCCGCTACAGGAAATTTGCAATTAAACGATCATCGAAAT 7608
QY      526 CAATTACGGAAT 536
Db      7607 TAACACGAGT 7597

RESULT 8
AB002583/c 149987 bp DNA circular pLN 07-MAY-2003
LOCUS      Cyanidioschyzon merolae chloroplast DNA, complete genome,
DEFINITION strain:10D.
ACCESSION  AB002583
VERSION     AB002583.1 GI:30409147
KEYWORDS    chloroplast Cyanidioschyzon merolae
            Cyanidioschyzon merolae
            Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.
ORGANISM    1
            Ohta,N., Matsuzaki,M., Misumi,O., Miyagishima,S., Nozaki,H.,
            Tanaka,K., Shin-i,T., Kohara,Y. and Kuroiwa,T.
            Complete Sequence and Analysis of the Plastid Genome of the
            Unicellular Red Alga Cyanidioschyzon merolae
            DNA Res. 10, 67-77 (2003)
            2 (bases 1 to 149987)
            Ohta,N., Matsuzaki,M., Tanaka,K., Shin-i,T., Kohara,Y. and
            Kuroiwa,T.
            Direct Submission
            Submitted (02-APR-1997) Nijii Ohta, Saitama University, Department
            of Molecular Biology and Biochemistry, 255 Shimo-ohkubo, Saitama,
            Saitama 338-8570, Japan (E-mail:nijiemolbio1.saitama-u.ac.jp,
            Tel:81-48-858-3848, Fax:81-48-858-3384)
            Location/Qualifiers
            1. 149987
            /organism="Cyanidioschyzon merolae"
            /organella="plastid:chloroplast"
            /mol_type="genomic DNA"
            /strain="10D"
            /db_xref="taxon:45157"
            1. 1398
            /gene="ycf24"
            1. 1398
            /gene="ycf24"
            /codon_start=1
            /transl_table=11
            /product="ABC transporter for iron-sulfur cluster
            formation (sucB)"
            /protein_id="BAC76097.1"
            /db_xref="GI:30409148"
            /translation="MNQCLNDMPQANREYAYGFSNPFDEPPRPIGNIENQIVMISAKK
            CEPQFPLDPRILKAYIHKKQMPRPQARVAYETINEQILTYAAPTSTEDVOSTIAKL
            GVAPEKVALDVPDVSJGTGHQALKDAGVTPCISIAIQKPSLVKXYASVPI
            GDNFSAIANSAYSDGSPCYIPQDTACPEMLSTYFRLQKSGSGQPERTLIIEKASV
            SYLEGCTAARYDHQHLAAVVELIALDHASIKYTVQNNYSGDQSGKGIIVYVRKG
            LCAEGAKISMTQVETGSAITWYPCISILAAINTADPFYSVALTDGYGQADGSMIH
            LGANSKRISIKSGICGGKSMNTYRGQVMTSAOKKAMNSQCDLLIGKATANTPY
            IETANSAVEHEAATTTOIGEBQLFTLQGRGINQDAVAVQMLTFCQGVFTQLDPRPEFA
            TEVNEILLTKBNC"

gene      1399..2139
            /gene="ycf16"
            1399..2139
            /gene="ycf16"
            /transl_table=1
            /codon_start=1
            /product="ATP-binding subunit of ABC transporter for
            iron-sulfur cluster formation (sucC)"
            /protein_id="BAC76098.1"
            /db_xref="GI:30409149"
            /translation="WKKIEQKVGIKDRLILEGIDAHIEAGSVHAINPNSGKSTLA
            KVIAPAPYAKLIEGNPFQDESINDKAEARAOQIFLAFQNPITTIQVNMDDLMA
            YKTHGKSIDAISFLIEYIMPELIVGIDASFLHRYVNDGSGGKKNEIILQALTEP
            KIALIDEISGLDIDALKOJASHLMLKTTQNAVYLITHYPRLLIYIKPDVHMVQKG
            RIISGASLAKYERYGYDEVKKAV"
            complement(2171..2275)
            /gene="ycf12"
            complement(2171..2275)
            /gene="ycf12"
            /codon_start=1
            /transl_table=1
            /product="unknown"
            /protein_id="BAC76099.1"
            /db_xref="GI:30409150"
            /translation="MNWHLAQLTVLAFVIAVGPIITLIMSKKDAL"
            complement(2193..3065)
            /gene="ycf62"
            complement(2193..3065)
            /gene="ycf62"
            /codon_start=1
            /transl_table=1
            /product="cell-cycle protein mesu family"
            /protein_id="BAC76100.1"
            /db_xref="GI:30409151"
            /translation="MYTPMHNKPNRROLANLTGSHSFSHSMILATISGQDSILKF
            MHDILHIVVHVDHDMSCSLARVRYLSRWHLAMRYFRTALPTITDQDAQVYF
            ANLQILQYHSHALCTAHRASDLELYLQWISLRPEGIWRLCHSGSCILFRPL
            HLTRAQTHFTLVIVLPYSDVSNYQTFPRNQRLQIFPLKXINFRFTTYSARLS
            ARILSPILVYFWHWSLPTLQYQATWGLTSLQIQLIYNNKVELARSSSINRS
            FCDSSRSYHFMAL"
            3086..3157
            /gene="trnK"
            3086..3157
            /gene="trnK"
            /product="trnK-Lys"
            /note="codon recognized: AAA"
            /anticodon="(pos:3118..3120,aa:Lys)"
            3183..3914
            /gene="trpA"
            3183..3914
            /gene="trpA"
            /codon_start=1
            /transl_table=1
            /product="tryptophan synthase alpha chain"
            /protein_id="BAC76101.1"
            /db_xref="GI:30409152"
            /translation="MLIAYLTGAPDINTTKEAVMKLAKKGADYIEIGVPSDALDG
            AILQKAKQALMNGFHDHMLNLSYNEIEVPLIATYINQIWHYGEKWKYKLVAH
            NKGILVDPDYESTKLRQICDRYGNIITLISPTNKRPADELARACKQWIVISR
            TGVLTLEEDKQIPKLIBELKVTXAPIALGFGISKSBOVKLVKSGADGVIIIGSAC
            MQLLEKGVQDLSEMIWVKKSF"
            3800..5140
            /gene="infB"
            3800..5140
            /gene="infB"
            /note="might be a partial remnant"
            /codon_start=1
            /transl_table=1
            /product="initiation factor 2"
            /protein_id="BAC76102.1"
            /db_xref="GI:30409153"
            /translation="MGGGRCNYKCHXANFIRKRSRSTIRMDKDEKILISVGTIVAD
            LAKQTKAPVTEFIRYAFVFTQGVIRLNDIVDEELMQLLQKRGWEMQINVEDETRADEVA
```

Query Match		34.5%;	Score 195;	DB 8;	Length 149987;
Best Local Similarity		65.0%;	Pred. No. 7.6e-44;		
Matches	319;	Conservative	0;	Mismatches	170;
				Indels	2;
				Gaps	2;
QY	48	CATGAAACACACCCTCTCTGTTTATTAGTTGAAGATGAAGCCGGAGTGTAAACCCGCAATTGC	107		
Db	11049	CATGAAACAAACTTATCTGTATATTAGTAGAAGATGAGCGAGGGCTTTTAATCTGAATTGC	1099		
QY	108	CGACATAATTTGCCCGCCGCTGTTTAACTTGAAGACTTGGCGTGGGCTCGCGGAGACA	167		
Db	10989	GGCGTTATTTGCCCGAGCTGACCTTTATATAGAAGTTTAGCAGTAGAGACTGTGAGAGCA	1093		
QY	168	GGGGGACGTTTCCCGCATCAACATGTGTGTGCCGGGGGATGAGACACCATCGAACACT	227		
Db	10929	AGCAGGTATATCTAGAAATTACCATGTGTGTTCCAGGAGACGATAGAACCATTTGAGCAACT	1087		
QY	228	GACCAACCAACTCTACAGTTGGTTAACGTAATTTAAAGTACAGGACATGACCCGAACTCC	287		
Db	10869	AATGAGCAATTAATATTAACGTGATACCCATTTTACAGTGGAAAATTGACACAGTACC	1081		
QY	288	CTGTGTGAAAGGGAATTGATGCTGTGTAAGTGAAGGCCCAATGCCCTTAACCGAGCGA	347		
Db	10809	TTGTGTGAAAGGAATTAATTTAATAGAAGGTGCAAGCAATCCACAGACAGAGAGAGA	1075		
QY	348	AGTGATAGACTAGCCAGGATATCCGGGGCCGCACTGTGTGATATCTCCGAAGAACAACCT	407		
Db	10749	AATTTTAGAATATGGCTAAATATTTTTTCCGCGACGTGTGTGATATATACAGACGATGT	1069		
QY	408	CACCATCGAA-TGTGTGGGAGACCCGGGTAAATAGTAGCAATCTCCAGATGTTGTGC-CA	465		
Db	10689	CATGTTAGAGTAAACGGGTATCTGTGAAAATATGATAGCACTGGAACAATTTTAAACGA	1063		
QY	466	AGTTGSCATTAAGAAGGTGGCTCGAACGGCCAAATTTGCTTTGGTGGGGAATCCGGCGT	525		
Db	10629	ATTGTGATATGTGGAAGTAGTACCCGATCAGGAAAAATTTGCATTAAACGATCATCGAAGT	1057		
QY	526	CAATAGCGAAT 536			
Db	10569	TAAACCGGAGT 10559			
RESULT 9					
LOCUS	PPU38804/c	191028 bp	DNA	circular	PLN 27-MAR-1998
DEFINITION	Porphyra purpurea chloroplast, complete genome.				
ACCESSION	U38804				
VERSION	U38804.1	GI:1276652			
KEYWORDS					
SOURCE					
ORGANISM					
		chloroplast Porphyra purpurea			
		Porphyra purpurea			
		Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;			
		Porphyra.			
REFERENCE		1 (bases 1 to 191028)			
AUTHORS		Reith,M.E. and Munmolland,J.			
TITLE		Complete nucleotide sequence of the Porphyra purpurea chloroplast			
		genome			
JOURNAL		Plant Mol. Biol. Rep. 13 (4), 333-335 (1995)			
REFERENCE		2 (bases 1 to 191028)			
AUTHORS		Reith,M.E.			
TITLE		Direct Submission			
JOURNAL		Submitted (17-OCT-1995) Michael E. Reith, Marine Biology Section,			
		NRC Institute for Marine Biosciences, 1411 Oxford Street, Halifax,			
		Nova Scotia B3H 3Z1, Canada			
FEATURES		Location/Qualifiers			
source		1..191028			
		/organism="Porphyra purpurea"			
		/organelle="Plastid;chloroplast"			

```
gene
CDS
/mol_type="genomic DNA"
/strain="Avonport"
/db_xref="taxon:2787"
181..1053
/gene="chlL"
181..1053
/gene="chlL"
/codon_start=1
/product="protochlorophyllide reductase iron-sulfur
ATP-binding protein"
/protein_id="AAC08073.1"
/db_xref="GI:1276653"
/translation="MKLAVYGGKGIKSTCSNISVALSKRKKYLQIGCPKIDSTF
TITGFLPITIDITLQAKDYHEDVWEDVITYGQGVDCVENGGPPAGACGGVYVGE
TVKLKELNAPDEYDILFDVLDGVVCGFAPLNVADYCLITDNGDALFAANRIA
ASVREKARTSLRLAGLVGNRTDKDLIDKYIDCVMPYLEVLPLEIDIRVSRVSKT
LFEMASDITNLAVCDYNYNIADQLIARPEGVVPKRAADRELFSLSDFYLNPKNTT
KSTADEEELDMV"
1095..2402
/gene="chlN"
1095..2402
/gene="chlN"
/codon_start=1
/product="protochlorophyllide reductase chlN chain"
/protein_id="AAC08074.1"
/db_xref="GI:1276654"
/translation="MSIVSDALTFECETGNYHTPCPISCVSWLYQKIEDSPFLVIGT
KTGCVFLQAMGVMIFAEPRVYMALEBEGDISAKNDYGLRLQLQIKDNPVIF
WIGCTTEIRIKMDLGIAIKLEAIEIRVIVARANGLDYAFQGBDYLAAACQPS
HHYIENTNHPVPLFEGSLDPVITQLWELVSGMIGKQIIVSGMLPSKRYELPIQGGY
VGVNPELSRTATTLTKRRRTKIGAPFIPGGPRAMTEIKISVSIKPEGLAREK
DWDSELYIALIRGKSVPRGMDNILEVLARFLRCGWTYVEIGIPVMDKYOAAEL
ALAKSTCDMNVMPTIVEKPDNYNOLDIRDLKPLVITGMANPPEARINTKMS
VEFTPAQIHGFVARDILSLVTRPLRNLISLELGDVYSKQ"
complement (2454..3929)
/gene="ORF491"
complement (2454..3929)
/gene="ORF491"
/codon_start=1
/protein_id="AAC08075.1"
/db_xref="GI:1276655"
/translation="NNFIQDLRLLLKSRYPILLINTREERLEYIIKQQLFCLNDEKY
YCMDFVGYTSNPNNGVAKRNPLIALEFIEKLDPEYALFLIKMDSPFNDIGIRG
LRMLSOIKTOSKNIIVGCKNVIPLYLSDVITVDLPNLSEIKETTRISLIGRI
VDLIELNNLAKSCOGSLVERIRKTIITKITAQGGDSRLPIIEKROLINORLU
EYRPYKKAIEDIGDALKSMLKRSRSKQSPNYGIPTRKGLLNGIQGSKSLTA
KATANDWTLPLKLDIGKLFGLVGBSSESRMEKVTIAGLSPCVIMDEIKASNL
YSGDSGTSARVFGPIITWSEKTPVFAVATANTQNPSEMLRGRDDELFDLDP
NMNERBLFIQIHILARIPKSMQNYNIKQSLCNKFSGAIEQAIYESMHTAFSEERE
FSTEDIKIALECFIPLAYTDKEQVESLQAWAGGRANASL"
complement (4047..4187)
/gene="trpJ34"
complement (4047..4187)
/gene="trpJ34"
/codon_start=1
/product="50S ribosomal protein L34"
/protein_id="AAC08076.1"
/db_xref="GI:1276656"
/translation="MTKQTIQSGSKRKIKHISGRPAKMTPSGRLINERRRGKKIM
AS"
4333..4854
/gene="ycf37"
4333..4854
/gene="ycf37"
/note="hypothetical chloroplast ORF 37."
/codon_start=1
/protein_id="AAC08077.1"
/db_xref="GI:1276657"
/translation="MLVTLPIFYLSITLFLTFSWLISQQLTILLESQRYFLDK
SKNDELIGDEGFASKVCIAKRYFTRAIIESHLVLRKNSLSESPENAAIADLYNMG
FIYKADQKLAKNFYERAIIEVDGNYIVALNNLAKIYEDTKNLIKAAALYDKVLIAK
SNKIANTHKEFIK"
complement (4889..4962)
tRNA
CDS
/gene="trnP"
complement (4889..4962)
/gene="trnP"
/product="tRNA-Pro"
complement (5020..5544)
/gene="ORF174"
complement (5020..5544)
/gene="ORF174"
/codon_start=1
/protein_id="AAC08078.1"
/db_xref="GI:1276658"
/translation="MIFMKTFPQNSTVGSYENKPKLLFDKTPDKRIELKQVLYSNT
KNITLYLEQDCDSYWKRPPLKKYIALKHSIIISLQKNSSTRRLVGFARATSN
GFNATIDVAVIHPPQGLGKVVVHQLIKQLQOAEISTITLFAEPVVISFYRKLGI
KDPDGVGMFWPP"
5807..6364
/gene="psaF"
5807..6364
/gene="psaF"
/codon_start=1
/transl_except="pos:5807..5809,aa:Met"
/product="Photosystem I reaction centre subunit III"
/protein_id="AAC08079.1"
/db_xref="GI:1276659"
/translation="MKKATFTICLAVLAVLSNPVIVNAEYAGLIPCKDSAAFNKRVN
SVKLOARIAKYDANPMPALAINKOIEKTRFATYGRAGLGGDGLPHLISGRMS
RAGDFVFPGLIFLYITTCMGVWGRGYLLSAKTSKPIRKETIILDVLAIVFMSSGFAM
PLAAWQEPSSQGLIAPDDITVSR"
6377..6505
/gene="psaJ"
6377..6505
/gene="psaJ"
/codon_start=1
/product="Photosystem I reaction centre subunit IX"
/protein_id="AAC08080.1"
/db_xref="GI:1276660"
/translation="MNNNFTKYLTAFVIGLVMTPTAGFIIELRFPFVLYYL"
6660..7145
/gene="apcd"
6660..7145
/gene="apcd"
/codon_start=1
/product="Allophycocyanin gamma chain"
/protein_id="AAC08081.1"
/db_xref="GI:1276661"
/translation="MSLVSOIINADDELRYPTIGELQSIONLYITGSNRLRIATIR
DKKEITIQKAKQIFQALPBYIARGNAAGRSRSLCPRYGWLRLIITYGVLAGDD
SIETIGIIGVREMTNSLGVPIIGMLDIQCLKEASLEMLGDDAKIISPFDYIIRGM
S"
7473..8453
/gene="fabH"
7473..8453
/gene="fabH"
/codon_start=1
/product="beta-ketoacyl-acyl carrier protein synthase III"
/protein_id="AAC08082.1"
/db_xref="GI:1276662"
/translation="MGVHILSTGSSVNFSEVNOQFDDIETSDHWISTRGIKRHL
APSSSTLRKLAABANDALSKASINADIDILIATSTPDLFESASQLQAEIATGS
TAPDITACSGFIITAYASQFIQAGSYNKVLVAGDPMRSMIDMSRSCILFQDGA
GAVLIGBSINSITGFLKCDGRNLSHQLMNSPSDQOQRLTTPPKRGYDSIIMNKG
EYKFAVQVPIVITKCLNDNISIDEDVMTILQANRIRILIAITLSLTSKITTN
LENYGNTSASISIPALDEAIKERTIQPQVVVLAGFAGLWGAIVLMQ"
complement (8465..8536)
/gene="trnA-Leu"
8735..8854
/gene="psbX"
8735..8854
/gene="psbX"
/codon_start=1
/product="Photosystem II protein X (4.1 kd)"
/protein_id="AAC08083.1"
/db_xref="GI:1276663"
```

	gene	/translation = "MTPSLSEFLNSLIGAVIVVPITLLALFVSCQDRITRS" 9115.. .9981 /gene = "accD" 9115.. .9981 /gene = "accD" /codon_start=1 /product = "acetyl-CoA carboxylase carboxytransferase beta subunit" /protein_id = "AAC08084.1" /db_xref="GI :1276664"
CDS		/translation = "MSLSNMPLEKKEGSAHNISTKTQTII PDGLMWKCPCDGILMYSK VLKNLKVCPCQSYHFQASNNERIDOLI DDOGSMEPMVHLISNDPLGFKOXKYSORL KDYLFPSLSDGAPVOEGMTLGQKVCGIIMDPFMGGSMGVSGEKLTRLEKATGER LPAILLCSSGARFMOEGMTLSLMQAKTISALEHKKENLLYSLVTSPTTGVTAFAM LMGGLLIABRPALLAIFAFRNRVIDETIEDLPDNQSSELYFEHGFDLLIVPRQLASK LIOLHLHTNNN" 10259.. .10750 /gene = "psbV" 10259.. .10750 /gene = "psbV" /codon_start=1 /product = "cytochrome c550 (oxygen-evolving complex component)" /protein_id = "AAC08085.1" /db_xref="GI :1276665" /translation = "MLKRSSWLTLLGLTVASVSITIVAIEIDEARTRVPLESNGRT VTLPEQGVRGRKRLFNNSCAICHNGGITTKNPVVGDPESLAGATPORDNIELVDYM KDLPTSDGAESIADIELHPRSIKSAEIFPQRNLUTDEDLFTIAHHILLPKIVEKWGCK IYY" 10915.. .11247
gene		
Query Match	31.8%; Score 179.6; DB 8; Length 191028; Best Local Similarity 62.7%; Pred. No. 1.8e-39;	
Matches 312; Conservative 0; Mismatches 184; Indels 2; Gaps 2;		
OY	49 ATGAACACACCCTCCTGCTTTGTGTGAAGATGAAAGCCGCAGTCGAACCCCATTTGCC	108
Dd	33511 ATGAACACTATTATCAAGCTCTAGTCCAAGAACAACGGAGCTTAATAAATAATCC	33455
OY	109 GGACTATTGGCCCCCGCTGATTTAAACATTGAGAGCTGGCGGTGGGTGCGCGAACA	168
Dd	33451 GGATTTATTGTGTAGCGGAGCTTTAAATATGCAAAGCTTGCTGTGTGACCGAAGAACA	33397
OY	169 GGGGACGTTTTCCGCATCACCATGTGTGTGTCGGGGGATGAAAACACATCGAACAATG	228
Dd	33391 ATTGAAGTTCCTAAATTAACGATGTGTGTGTCGAAGACATTAATGAACAATTGAGCAACTA	33332
OY	229 ACCAAGCAACCTCAAGTAGTGTTAAGCTAATTAAGATACAGACATCACCGAACTCCC	288
Dd	33331 ACTAAACATTAATAATTAGTTAATCATCTTNATATCAAGATGAACCAATTAATCCT	33272
OY	289 TGTGTGAAAAAGGAAATTGATGCTGTGTGTAAGTAGACGCCCAATGCCCTTAACCGAGCGGA	348
Dd	33271 TCCTGTGAGAGAAATTAAATGTATAATTAATAATTCAAGTTAACTACCAAAACAATAAGAA	33212
OY	349 GTGATTGAGAGCCCGAGTATTCGGGGCCCGCATTTGTGATATTCGCCAAGACACCGTC	408
Dd	33211 GCTCTGGAATTTGMAAAATTTTTCAGACCAATGTAGTAATATGCGGAAGATTTAATA	33152
OY	409 ACCATCGAA-TGGTGCGGGAACCCGGGTAAATGATGACATCCTCCAGATGTTGACCAAG	467
Dd	33151 AATTGTGAAGTAATCGAAGACCTCGAAAAAATGTGGCATTAAGAACGATATTGACTAAA	33092
OY	468 TT-GGCAATTAAAGAGGTGCTCGAAGCGGCAAAATTGCTTTGTGTGGGAAATCCGGCGTC	526
Dd	33091 TTCGGCAATTATGAAATAGCTAGAACGTGAAGAAATTTGCTAGTAAGAACATCAAAAATT	33032
OY	527 AATACGAATATCTGAAA 544	
Dd	33031 AATACGAATATTTAAAA 33014	

AF233069/c	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
AF233069	6982 bp	DNA	linear	PLN 04-JUL-2000								
Galdieria sulphuraria	matutase (matK) gene, partial cds; 50S											
ribosomal protein, ribulose-1,5-bisphosphate carboxylase/oxygenase												
large subunit (lbcu), ribulose-1,5-bisphosphate												
carboxylase/oxygenase small subunit (lbcS), acetylhydroxy-acid												
synthase, small subunit (lbcS), and translation initiation factor												
IF-3 (lbcS) genes, complete cds, and unknown genes; chloroplast												
genes for chloroplast products.												
AF233069												
AF233069.1	GI:8925945											
chloroplast Galdieria sulphuraria												
Galdieria sulphuraria												
Eukariota; Rhodophyta; Bangiophyceae; Porphyridiales;												
Porphyridiaceae; Galdieria.												
1 (bases 1 to 6982)												
Whitney, S.M. and Andrews, J.												
Direct Submission												
Submitted (08-FEB-2000)												
Molecular Plant Physiology, Research School												
of Biological Sciences, P.O. Box 475, Canberra City, ACT 2601,												
Australia												
Location/Qualifiers												
1. .6982												
/organism="Galdieria sulphuraria"												
/organella="plastid;chloroplast"												
/mol_type="genomic DNA"												
/strain="UTEX 2393"												
/db_xref="taxon:130081"												
complement (<1. .407)												
/gene="matK"												
complement (<1. .407)												
/gene="matK"												
/codon_start=1												
/product="matutase"												
/protein_id="AF81678.1"												
/db_xref="GI:8925946"												
/translation="MKHKOSIFPKKFKLFFKYLWNIHPSLGARSVPIEMSYKYNIA												
RAILNLYLVKXYSVYALAKEDYINKILNRNQLAERYAIDLETYDLITAKEX												
REAYDNMFYRKYYLFIDRLRKGHITFGGFDL"												
complement (522. .860)												
/note="ORF39; hypothetical 12.5 kDa protein"												
/codon_start=1												
/product="unknown"												
/protein_id="AF81679.1"												
/db_xref="GI:8925947"												
/translation="NMTYINIRKSLDLHNSISKLSNEDKLYIRIGIKGGCGSLSY												
FMVYKKSNTSDNDIVYDNPFLVCDNKSILYLGLSDYSSLLDGRKFLPNPKK												

GMHOLIVHIGEDVYLQEGGCTTGHDPDIOGATANRVALLAMILLARNENDPFLTEGP
ELIREAKXCGALRTALDLMKDITFRYISTDTSDFVEPTTANI"
gene 2859..3278
/gene="tbcS"
CDS 2862..3278
/gene="rbcS"
/codon_start=1
/transl_except=(pos:2862..2864,aa:Met)
/product="ribulose-1,5-bisphosphate carboxylase/oxygenase
small subunit"
/protein_id="AAF81682.1"
/db_xref="GI:8925950"
/translation="MRITQGTSPFLPDLTDEQIKKQIDYMISKLAIQIETNDIHPR
NSFMWMLPLFEVNDPAPVLPFINACRAKSNFYIKVGFSESGISTTISFVNR
PKHEPGRNLIRQEDKRSIKSISQAVETYPEDQRY"
3391..4308
/note="similar to CfxQ/CfxX protein"
/codon_start=1
/product="unknown"
/protein_id="AAF81683.1"
/db_xref="GI:8925951"
/translation="MVVITPEYKEENIMNTLVNIOEYKTKIQEVINOLDKELIG
LTPVKRRIKELALILIDKRIKLPQANPGLHNSFTSPGTCTTATKADLIYL
LGYIRKGLLTVTRDVLVQYIGHTAPKREVLKRAMGVLDIDAVYLYKPDNRDY
GSAIILILQWENQNRDLVILAGYKDMRFYQSNPLSRITNHDVPDYAEEL
LEIGIMLEEQOYLTEGKVLLEYIKRMTQPHANARSIRNAIDRAMQOANRIF
NSSNRYLTKKDITLIEADILKSLFMNR"
complement(4506..5051)
/gene="ahaS"
complement(4506..5051)
/gene="ahaS"
CDS 5201..5746
/complement(4506..5051)
/gene="ahaS"
/codon_start=1
/product="acetylhydroxy-acid synthase small subunit"
/protein_id="AAF81684.1"
/db_xref="GI:8925952"
/translation="MKHTLSVVEDSSGLTRISGFARGENIDSLANGPRAKGIS
RITWIRGDRRIIEQITKHLKLVNLKIQDITNIPCVREMLIKVKTGENKREIL
EIANIFRARVLDISEFIILETIGDPKLAIEKLEKYGICEIATGRIALTRRESKI
NTEILRTWIMNDKIRGGGT"
5201..5746
/gene="1f3"
5201..5746
/gene="1f3"
/codon_start=1
/transl_except=(pos:5201..5203,aa:Met)
/product="translation initiation factor IF-3"
/protein_id="AAF81685.1"
/db_xref="GI:8925953"
/translation="MLKSKQKNSPDKNKFIINNKISFPIVRIIDENVOIGYKIED
AKLAEKNDLVLINDKIBPVRIIDGKFPRAEKSEKROHYVTEIKR
KIEDHDYQRIQALIKFLGNKVKISLTFKREIQYIDLAKLLEKLIHSLSLAE
HEKVPDEKGNIIIMCLPKKT"
5786..6505
/note="ORF720: putative ABC transporter; ATP-binding
protein"
/codon_start=1
/product="unknown"
/protein_id="AAF81686.1"
/db_xref="GI:8925954"
/translation="MKILIEKCYFARKCILDNLHMTSQVEILGLPNGTKTISI
FKLTIGSKDKGVLNGDITNPMHQARLGIQYLPSSVFRLLTVKFKLAL
EMGYSAPASISITKSTFYQIDKTLHFGYSIGSEERKKEIAAVIIGKFIILD
EPAGIDPISISLQITIVNLNNSGICILITDHWVETLKITHRSYILSGTNILAGT
TRQILNKAVRESLGNFTL"
6553..6981
/note="ORF429"
/codon_start=1
/product="unknown"
/protein_id="AAF81687.1"
/db_xref="GI:8925955"
/translation="MFTITRILIQLLIGFVSSAFTSITLSVGLPELVARVVEG
LPEKIALKILVKSPTIFLSPPLASFAVLITFNQIANKKELIGFRSGISNFKWL
STIFGTILAISSFINDLIIVPSNHRANELLRESKENTL"

ORIGIN

Query Match 31.3%; Score 177; DB 8; Length 6982;
Best Local Similarity 62.8%; Pred. No. 7,56-39;
Matches 307; Conservative 0; Mismatches 180; Indels 2; Gaps 2;
QY 49 ATGAACACACCCCTCTGTTTGTAGTTGAAGATGAACCCGAGTGCTAACCCGATTGCC 108
DB 5051 ATGAACATCATATATCTGTTTGTAGTAGAAGATGAACCCGATTGAATCCGATTTC 4992
QY 109 GGACTATTGTCGCCCGCTGTTTGAATTGAGAGCTTGGCGGTGGCGGACAG 168
DB 4991 GGATTAATTGTTACACGCGATTCAATATAGTATAGCTTACAGATTGCTCTGCAAAA 4932
QY 169 GGGGACGTTTCCCATCACCATGATGATGATGATGATGATGATGATGATGATGATGATG 228
DB 4931 GCAAGTATATCAAAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATG 4872
QY 229 ACCAAGCACTCTCAAGTGTGTTACGTAATTAAGTAAGACATCACCAATCTGCC 288
DB 4871 ACAAACATCTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4812
QY 289 TGTGTGAAGAGGAATTGATCTGTGTAAGGTGAGGCCCAATGCCCTAACCGCGGAA 348
DB 4811 TGTGTGAAGAGGAATTGATCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4752
QY 349 GTGATTGAGTAGCCAGTATTCGCGGCCGATTTGATGATATCTCCGAGACACCGTC 408
DB 4751 ATATTGGAATATGCAATATTTTATGAGCACTGTGTTGATTTATGAAATCTTTATT 4692
QY 409 ACCATCGA-ATGCTGGGGGACCCGGTAAATGTAGCAATCTCCAGATTGGCCAG 467
DB 4691 ATTTAGGATTAACAGGATCTGCGCAAAATCTGCTATGAAAAATTATTAAGAAAA 4632
QY 468 T-TGGCATTAAGAAGTGGCTCGAACGGGCAAAATTTGTTTGTGGGAAATCCGGCGTC 526
DB 4631 TATGTATTGTTGAAATGCTTGAACAGCAAGATGCTTGAAGAGATCAAAAATA 4572
QY 527 AATACGAA 535
DB 4571 AATACTGAA 4563

RESULT 11

EX572098 346683 bp DNA circular BCT 23-SEP-2003
LOCUS Prochlorococcus marinus MIT9313 complete genome; segment 4/7.
DEFINITION BX572098 BX548175
ACCESSION BX572098.1 GI:33640689
VERSION
KEYWORDS complete genome.
SOURCE Prochlorococcus marinus str. MIT 9313
ORGANISM Prochlorococcus marinus str. MIT 9313
Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
Prochlorococcus.

REFERENCE

1 (bases 1 to 346683)
Rocap,G., Larimer,F.W., Lamerdin,J., Malfatti,S., Chain,P.,
Ahlgren,N.A., Arellano,A., Coleman,A., Coleman,L., Hauser,L., Hees,W.R.,
Johnson,Z.I., Land,M., Lindell,D., Post,A.F., Regala,W., Shah,M.,
Shaw,S.L., Steglich,C., Sullivan,M.B., Ting,C.S., Tolonen,A.,
Webb,E.A., Zinser,E.R. and Chisholm,S.W.
Genome divergence in two Prochlorococcus ecotypes reflects oceanic
niche differentiation

TITLE

Nature 424 (6952), 1042-1047 (2003)

JOURNAL

MEDLINE

PUBMED

12917642
2 (bases 1 to 346683)
Larimer,F. and Rocap,G.
Prochlorococcus genome consortium
CONSRM
Direct Submission
TITLE
Submitted (03-JUN-2003) Submitted on behalf of the Prochlorococcus
genome consortium, the DOE Joint Genome Institute, Production
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,

1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov
location/Qualifiers
1..346683
/organism="Prochlorococcus marinus str. MIT 9313"
/mol_type="genomic DNA"
/strain="MIT9313"
/db_xref="taxon:74547"
complement(347..637)
/locus_tag="PMT0970"
complement(347..637)
/locus_tag="PMT0970"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Herpesvirus VP23 like capsid protein"
/protein_id="CAE21145.1"
/db_xref="GI:33640690"
/translation="MRKTHRIHRIFEGTPEPLLCIRKSKOTLLSCDLVDIEPLLA
WIAFGAVDLIFSNKRPDILLRLGATFIVLALSHVSPSDVCKASIMQHALS"
620..1015
/locus_tag="PMT0971"
620..1015
/locus_tag="PMT0971"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible bZIP transcription factor"
/protein_id="CAE21146.1"
/db_xref="GI:33640691"
/translation="WDFDSQKPTSRSGDYYRPAVLTITLVRSSEILMASCGRGLE
INDYSRPEBYNGPFGITIPDRVYVDPDSYKKAQRONQRYARLERRARELTLY
VDLEKQIMELQPSQDEVPDKGLAAKT"
1478..1699
/locus_tag="PMT0972"
1478..1699
/locus_tag="PMT0972"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible DNA polymerase (viral) N-terminal dom"
/protein_id="CAE21147.1"
/db_xref="GI:33640692"
/translation="WPNQRLRRLLEFLKPRVLAAGSEFFTPMQSKAGIDCITLRW
LSDVPEALALDDQLKQVLDQARWLYN"
2436..2690
/locus_tag="PMT0973"
2436..2690
/locus_tag="PMT0973"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Vinculin family"
/protein_id="CAE21148.1"
/db_xref="GI:33640693"
/translation="MQGFMYGDERMLHRAVALFSCFRWLGWTMGLFSKRDLQVEFKP
PKEVTFKPDLDLSPISRDEPFKRLLEQTLAESPSQED"
complement(2865..3230)
/locus_tag="PMT0974"
complement(2865..3230)
/locus_tag="PMT0974"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAE21149.1"
/db_xref="GI:33640694"
/translation="WQMDRNVLGEPLIEICSDPYTGMVYRQCKRTDGHDLGQHTVCCV
ISEQFLYISKAQNDLSTPIQYGFSGIKPDHWCVCASRWKQAYDDGMAPNIRLEXT
NIEVLAIDILKQYSHT"
complement(3351..4007)
/locus_tag="PMT0975"
complement(3351..4007)
gene
CDS

sig_peptide
gene
CDS
misc_feature
gene
CDS
/locus_tag="PMT0975"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Ribosomal protein L36"
/protein_id="CAE21150.1"
/db_xref="GI:33640695"
/translation="MRKPSIDRIFSKALAGILLCTSTISAEXKDYIEGTTYSNCR
EGIMKSKTIGRGVNDKEMTITERLKLSTGRQVNRNDIRKSCQSSLLSKARATGA
DCELASDFPSNWDIYASRFNAHGNKHITFAIRYDGSRLCLSKNNFRDLPLET
PGRSLFLNHERQADSSIIIEYEHYPNGWGYEVKKYRLDLSPEPSPFSVDSWIIHOR
"
complement(3918..4007)
/locus_tag="PMT0975"
/note="Signal predicted by SignalP 2.0 HMM (Signal peptide
probability 1.000) with cleavage site probability 0.694 at
residue 30"
4938..5267
/locus_tag="PMT0976"
4938..5267
/locus_tag="PMT0976"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Protein of unknown function DUF42"
/protein_id="CAE21151.1"
/db_xref="GI:33640696"
/translation="WGFQALQVGRINQAGIDAPLPIAVGENSVFIDALYANPSF
YDNVSSIIINTVPMKACPFGRKEGTMLPEFYKGVSGMGEILLEDDQSRVVSQILISAG
PAAVND"
complement(5257..7404)
/locus_tag="PMT0977"
complement(5257..7404)
/locus_tag="PMT0977"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="ABC transporter, multidrug efflux family"
/protein_id="CAE21152.1"
/db_xref="GI:33640697"
/translation="MTTNDPGRALIQARALEELLPSYPENEGEIEELRLGFCVAKL
GRQKLOPLSGASIKITLNFNDIHRKIEAPLDNQOAEPLVYVDFAEHNPFLAYR
HGGQWMLPDDSESHRLQDSNKLKLAATAEVYPSIPVYGLAIVRAPARLCAIT
ALVITTAAYMVLFTLSLPMLTNVVSRILPQNDQNLFGQTLVAVITIVGSAETELQ
LMLRLSTIADLRLOSAVNDRLRLPMSFTISXYTGGDLASRSALISQRLQNGEVL
TLLSLFAISFVLMEVYDNOALWASATLVAALCLMTITVRSIQLORPLESGAEI
TNLSLQAVMGPOIIRSGSEPYLVWRLREIRYASLQRSVYSDAMEQYGLVTP
ASLPMFAVVAOVVNSPSEINOTLVAIFSNAAFGNGALTGAINLIANVAGRAA
VLMRREPVLVADVEKGYQPEALVHSLGSEFRLRGISYEPPSSROLNLSLTIPAG
OHTAITGASGCGKTLVYMLTGFTISPGCEILVYDGIPTQALIRAYRQLVGMOTAR
LNTGSIVYICGCGVRSBEVQALPQALADRYAMPQVLETLSDSGGINSQGOV
RIARALITQPRVLIIMDEATSLDNRSQERTQTIKEIGMTRISIAHRLSTIQOAO
IYLERNQPEVWSGKMDLKGHYLASLASH"
order(complement(6868..6933), complement(6757..6822),
complement(6523..6588), complement(6451..6507),
complement(6187..6252), complement(6076..6111))
/locus_tag="PMT0977"
/note="6 probable transmembrane helices predicted by
TMHMM2.0"
complement(7401..9617)
/locus_tag="PMT0978"
complement(7401..9617)
/locus_tag="PMT0978"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="ABC transporter, multidrug efflux family"
/protein_id="CAE21153.1"
/db_xref="GI:33640698"
/translation="MKLQKRVKQAQRLQYAEACGASALATILAYRGLVEISELR
IACGVNRGSNAKQVLAARVGLAKYVRSQSGTBSGGRPCIVFQGNRFLVYRG
RNDRAVPLSDPAQGRIVMEBFDDKFTGIYVLEBPBPGRPGESRSPDLSLPAKLL
PGRSLDQLCVASQGVFGIIVAGFTSGFTIGLQNRRLYFGIPIIWLMLSLVGLWL

VERSION	AF022186.2	GI:6466296
KEYWORDS		
SOURCE		
ORGANISM	chloroplast <i>Cyanidium caldarium</i>	
REFERENCE	<i>Cyanidium caldarium</i>	
AUTHORS	Eukaryote; Rhodophyta; Bangiophyceae; Porphyridiales, Porphyridiaceae; Cyanidium.	
TITLE	1 (bases 130696 to 132364) Vogel, H., Fischer, S. and Valentín, K. A model for the evolution of the plastid sec apparatus inferred from secY gene phylogeny	
JOURNAL	from secY gene phylogeny	
MEDLINE	Plant Mol. Biol. 32 (4), 685-692 (1996)	
PUBMED	97134960 8980520	
REFERENCE	2 (bases 1 to 164921) Glockner, G., Rosenthal, A. and Valentín, K.	
AUTHORS	The structure and gene repertoire of an ancient red algal plastid genome	
TITLE	J. Mol. Evol. 51 (4), 382-390 (2000)	
JOURNAL		
MEDLINE	20496959	
PUBMED	11040290	
REFERENCE	3 (bases 46857 to 47851) Valentín, K.	
AUTHORS	Direct Submission	
TITLE	Submitted (22-MR-1996) Institute for Plant Physiology, Justus Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany	
JOURNAL	4 (bases 28701 to 73580)	
REFERENCE		

TITLE	Direct Submission
JOURNAL	Submitted (02-SEP-1997) Department of Genome Analysis, IMB Jena, Beutenbergstr 11, Jena 07745, Germany
REFERENCE	5 (bases 1 to 164921)
AUTHORS	Gloeckner,G., Rosenthal,A. and Valentin,K.
JOURNAL	Direct Submission
REFERENCE	Submitted (18-NOV-1999) Genome Analysis, Institute for Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
AUTHORS	6 (bases 130696 to 132364)
TITLE	Vogel,H., Fischer,S. and Valentin,K.
JOURNAL	Direct Submission
COMMENT	Submitted (18-NOV-1999) Institute for Plant Physiology, Justus Liebig University, Heinrich Buff Ring 58-62, Giessen 35392, Germany On or before Nov 23, 1999 this sequence version replaced gi:529651, gi:1240002, gi:2465730.
FEATURES	Location/Qualifiers
source	1..164921 /organism="Cyanidium caldarium" /organella="plastid:chloroplast" /mol_type="genomic DNA" /strain="Rkl" /db_xref="taxon:3771" join(164808..164921,1..78,462..644) /note="direct repeat; repeat unit separated by stem loop" 339..477 /note="stem-loop separates direct repeat unit" 580..780 /note="similar to yeast mitochondrial origin of replication region in Genbank Accession Number U36902" 986..2041 /gene="psbd" 986..2041 /gene="psbd" /product="unknown" 986..2041 /gene="psbd" /note="Photosystem II D2 protein" /codon_start=1 /transl_table=1 /product="unknown" /protein_id="AAPI3026.1" /db_xref="gi:6466445" /translation="MTIATIGRGEORGEWFDLLDMKIDRFVFIQSGTILFPCAYLAL GAMTGTTFVSSWTHTGLASGSCNPFTLPAVSSPANSNGSHLLPLMGPAQDPTFR WCQGGIMTFTALHGSFOLIGFCRQGEPIETALVLRPNALAFSGPIAIVFVSLLIVR
CDS	

gene
mRNA
CDS
/gene="ycf19"
/product="unknown"
/protein_id="AAFI3025.1"
/db_xref="GI:6466444"
/translation="MLFNKDKNNIGRSIESTGFAMSGNARLINTLSGKLIGAHANA
GLIVFTGAMTLFETSHFPEKPLVEQGMILLPHLATLGMVAPGGEIVNTYFEAG
VHLVSAVAGGAGIYHSIVGPDVLEDSFSPFGYWRDKNMTTILIGLILGAF
LIVTALFICGIVDTMAPGGDIRITNTLPALIPSYLKSPPGCGMTVGWNE
DVIGKHWIGTVCVIGIWHITRPFNSNARAFTVSGEAYISYSLGALALMGOTAAEY
AMTNTVTPSEFTGTAABASQAQFTLVKDQRLGANIATGQGTGKYLMSPTG
EVLIGETWRFWDLRAPWLEPLRSSNGLDLNKIKNDIQWERRAAEYMTAPLGSIN
SVGVATEINSYVSPRSWLTTSHPFLGFIPLGHLWHAGRAAAAAGEFKGINREN
EVLMSRPID"
3546..3821
/gene="ycf19"
/note="similar to ycf19 in Porphyra purpurea"
3546..3821
/gene="ycf19"
/product="unknown"
3546..3821
/gene="ycf19"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAFI3024.1"
/db_xref="GI:6466443"
/translation="MKDGTDLLETTIAFLQIYVILLRSLGMPFNINWSSQFYS
LSQSDPYLNFHGVSPFLGIDFSPILGILIDPIIELLSRQLKPL"
3873..4133
/gene="ccrf9"
3873..4133
/gene="ccrf9"
/product="unknown"
3873..4133
/gene="ccrf9"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAFI3023.1"
/db_xref="GI:6466442"
/translation="MNNCLPKNKVFKILNPFKCIDPMCTDNLKFPPLKLNKNGKFT
NKFYNTKNKKYIDWLFIYIYHILRIGNQIKDQITKC"
4189..4272
/gene="ccrf10"
4189..4272
/gene="ccrf10"
/product="unknown"
4189..4272
/gene="ccrf10"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAFI3022.1"
/db_xref="GI:6466441"
/translation="MTKLKSLNDKLTLLISLNTKIYIDIN"
4334..4406
/product="rRNA-Thr"
4450..4971
/gene="ycf3"
/note="similar to ycf3 in Porphyra purpurea"

mRNA
CDS
4450..4971
/gene="ycf3"
/product="unknown"
4450..4971
/gene="ycf3"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAFI3021.1"
/db_xref="GI:6466440"
/translation="MPSQKNDNFIDKTESIIADILVKILPTNKSEKAFYYKQDMA
AGESEGYAEKLCYQALKEIDPMKSPILYNIGILOKNGQHRALEYHESIKFN
PNLVQALNNAVIVHYGKLPFGSLQKAKMFPKASNYKALKAPYNTIEQNW
LKTTGRITEDIML"
5004..6476
/gene="atpB"
5004..6476
/gene="atpB"
5004..6476
/gene="atpB"
/product="unknown"
5004..6476
/gene="atpB"
/note="Atp synthase CFI beta chain"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAFI3020.1"
/db_xref="GI:6466439"
/translation="MKPCQKKOKQTLIAKYNKKGKVIQIIGPVLIVFPDQLPKV
FNAIKTNNSNNWITCEVOQLGDNKVRPAVSTTEGLKGSALIDTSPISIPGKE
TLGRIFNVLGSEPIDEKVPSINDKLPIDHRAPEKFOLEKPSIFETGIVKVDLAPYR
RGGKIGLPGAGVGVKIVLIMELINNVAKAHGVSVFGVGERTREGNDLYOMKESGV
INEKDNLTSKVALCYGQMBPPGAMRVGLTALTAAYFRDVNKNVLLFIDNIPRV
QAGSEVSALLGMPBSAVGYPGLTGEMLQORITSLDGSITSIOAYVAPADLIDTP
APATTPAHLDATTVSRALAAKGIYPADVLDSTSMLOPGIVSEHYTARKVETL
QYKXELQDITAILIGDELDSEEDRLIVSRARKIEKLSQPPFAEVFTGISGKTVSLD
SIRGFNMILSGEVDNIPGQAFYLVGRIEBALDKAQKERS"
6486..6887
/gene="atpE"
6486..6887
/gene="atpE"
/product="unknown"
6486..6887
/gene="atpE"
/note="Atp synthase CFI epsilon chain"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAFI3019.1"
/db_xref="GI:6466438"
/translation="MALTVKVTTPDRVVKKTVEIIIPSSTGQGLIMNHAPLITL
DIGWRARWNTWVPLVLDGFAQIDNNLVITIIYSDAEVXAIDEENKULLAASLAN
Query Match 30.7%; Score 173.2; DB 8; Length 164921;
Best Local Similarity 61.8%; Pred. No. 1,2e-37;
Matches 308; Conservative 0; Mismatches 186; Indels 2; Gaps 2;
QY 49 ATGAAACACACCCCTCTCTGTTTAAAGTAGAAGCCGAGTCTAACCCGCAATGCC 108
DB 106832 ATGAAATATACCTCTCCCTATAGTAGAAGTAGAAGCGGCTTTACACGTATCCCA 106773
QY 109 GGACATATTTCCCGCGCGGCTTTAAATGACATGAGACTTGGCGGCGGAGAG 168
DB 106772 GGCTTGTTCCTAGACGAGGATTTAACATAGAAAGTTTACGTGCTGCTGACGAA 106713
QY 169 GGGGAGCTTCCCGCATCCATCGTGTGTCGCGGGGAGTAGAACACCATGCAACACTG 228
DB 106712 AAAGCATCTCAAGATTCATGATGTTGACCTGGTGTATGACGACAGTTGAACAGTTA 106653
QY 229 ACCAAGCACTCTAACAAGTGTGTTAAGTAATTAAGTACAGACATGACCGAACTGCC 288
DB 106652 ACAAGCACTTTAATTAAGTATATTTTAAAGTAGATATATTAATTAACAGATTCT 106593
QY 289 TGTGTGAAAGGAAATGATGCTGTGTAAGTAGAGCCCAATGCCCTTAACCGAGCGGAA 348

Db 106592 TGTGTGAGCGTGAACACTAGTACTAGTAAATAATGCTAGTACTAGTCCAGACACAGA 106533
Qy 349 GGTATGAGCTAGCCAGGATATTCGGGCGCCGATGTGGATATCTCCGAACACACCGTC 408
Db 106532 ATTTTAGTATTTTGGCAGATTTTAAAGTGAAGTGTGATTTTACAGAAAATCTGTTA 106473
Qy 409 ACCATCGAA-TGGTGGGGACCCGGGTAAATGTAGCAATCCTCCAGATGTTG-GCCNA 466
Db 106472 GTTTTAGAATGTGACAGGGGATCCAGGTAAAGTAGCTGCAATTCACAGCTACTGCAAAA 106413
Qy 467 GTTGGCATTTAAGAGGTGGCTCGAAGCGGCAAAATTGCTTTGGTGGCGGAATCCGGCGTC 526
Db 106412 TTGGATATATAGAAATTCAGAAACTGGCAAAATAGCGTTAACAGAACATCTAAAGTA 106353
Qy 527 AATACGGAATATCTGAAA 544
Db 106352 AACACAGAGTTTGTGAAA 106335

RESULT 13
AE017164 302399 bp DNA linear BCT 14-AUG-2003
LOCUS Prochlorococcus marinus subsp. marinus str. CCMPI375 section 4 of 6
DEFINITION of the complete genome.
ACCESSION AE017164 AE017126
VERSION AE017164.1 GI:33237969

KEYWORDS Prochlorococcus marinus subsp. marinus str. CCMPI375
SOURCE Prochlorococcus marinus subsp. marinus str.
ORGANISM Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
Prochlorococcus.

REFERENCE 1 (bases 1 to 302399)
Dufresne,A., Salanoubat,M., Partensky,F., Artiguenave,F.,
Axmann,I.M., Barbe,V., Duprat,S., Galperin,M.Y., Koonin,E.V., Le
Gall,F., Makarova,K.S., Ostrowski,M., Oztas,S., Robert,C.,
Rogozin,I.B., Scanlan,D.J., Tandeau de Marsac,N., Weissenbach,J.,
Winker,P., Wolf,Y.I. and Hess,W.R.
Genome sequence of the cyanobacterium Prochlorococcus marinus
SS120, a near minimal oxyphototrophic genome
Proc. Natl. Acad. Sci. U.S.A. 100 (17), 10020-10025 (2003)
2 (bases 1 to 302399)

JOURNAL Dufresne,A., Salanoubat,M., Partensky,F., Artiguenave,F.,
Axmann,I.M., Barbe,V., Duprat,S., Galperin,M.Y., Koonin,E.V., Le
Gall,F., Makarova,K.S., Ostrowski,M., Oztas,S., Robert,C.,
Rogozin,I.B., Scanlan,D.J., Tandeau de Marsac,N., Weissenbach,J.,
Winker,P., Wolf,Y.I. and Hess,W.R.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2003) UMR 7127 CNRS et Universite Pierre et Marie
Curie, Station Biologique de Roscoff, Place Georges Tessier BP74,
Roscoff 29682, France

FEATURES
source location/Qualifiers
1..302399

organism="Prochlorococcus marinus subsp. marinus str.
CCMPI375"
mol_type="genomic DNA"
strain="CCMPI375; SS120"
sub_species="marinus"
/db_xref="taxon:167539"
103..1146
/gene="acsf"
/locus_tag="Pro0992"
103..1146
/gene="acsf"
/locus_tag="Pro0992"
/note="PM: 10611605; di-iron enzyme CcdI expression is
activated in copper- or oxygen-deficient cells"
/codon_start=1
/transl_table=11
/product="Mg-protoporphyryn IX monomethylster aerobic
cyclization protein homolog"
/protein_id="AA000037.1"
/db_xref="GI:33237970"

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

/translation="MTTTS GAPASLGRNELPHLDENLTPRPYTTPEKAAKTDL
IARDPEAMFEKEMADYNLKHPRKASLERLDLSPEDKAVESYLVRVSEFSGFL
LPRKISRPFKAKROELGCFPLIARBEAHAGLGLAALTREGINDLPHLPKRAT
FFPLSVTLVLSYEKTLGWRYLLINHLAANDKVCAPLDFPEPCQDENHGDCT
NLMWRCPWMTKGRKGLSRFLMTVFLTLTVCRGSEFYELGIDPLDPEVLI
QTNNTSNAPWPVYKPEDGKFLSRIDILNAFKMRQNGIKKPLAGKFLVLLKQF
TLPEKTDVARYG
complement (1153..1515)
/gene="petF"
/locus_tag="Pro0993"
complement (1153..1515)
/gene="petF"
/locus_tag="Pro0993"
/codon_start=1
/transl_table=11
/product="Ferredoxin, petF"
/protein_id="AA000038.1"
/db_xref="GI:33237971"
/translation="MKSHKITVLRDQNRILSKIEPESGILSKFEEDGVLPPSCRN
GCCITCAVRILSGTIDQHDGIGSKQWENGVLCTIAKVTGADLETQDESYELQ
FGRYLGSIKRRTGNPFDM"
complement (1538..1870)
/locus_tag="Pro0994"
complement (1538..1870)
/locus_tag="Pro0994"
/codon_start=1
/transl_table=11
/product="Predicted protein"
/protein_id="AA000039.1"
/db_xref="GI:33237972"
/translation="MTDQCKLYVLAPIGQLCGNGQLRRTISRRNRKGDVAFWYLS
ADLVKRLNINLSTFEAVVADKTAIDMLKLPFGELSLINDIDLEILKQDAMDLPAP
IKDISIPD"
1999..3450
/gene="tliD"
/locus_tag="Pro0995"
1999..3450
/gene="tliD"
/locus_tag="Pro0995"
/note="tliD"
/codon_start=1
/transl_table=11
/product="Predicted Zn-dependent protease"
/protein_id="AA000040.1"
/db_xref="GI:33237973"
/translation="MKVALTKAFSSNPLKTDQIDHLKRNNEELLAYGISAGADVE
IFENLDSIAFLAEDPTVTSVNPISGKGVIRFLANKGPFVSTNPSGDLFPALDQ
ALMLGILERNISIKSYFNGLINVDYKAKSGYLSCLPSIKESTNKLKLTSLQFG
PHLSVRGYSKTFGEVIVASDGTGADRLRTLSIGLVLAIDKGRYSIGRISA
GKNPTLLDMABEATAKINEASKMLYAVYVSGWPAVLNARGGVIRPHEAGCHLE
TQIERTGTPEBKANQOIAKNAVTVVDEGLITDCAFGSISMDDEGKETQTVITNGI
LKQFLSPRAGELEKTHNRKTSGRQSYPAPAAKRMNTYIEKSPVQPNELISIDNGI
YCKSGGSGVAGATGQNFSEBGLIKNGELTDPVGAATLIGBAKEMPRISMCAIDL
DLAAGCGSISGINVTGQPHIKVDSITVGK"
3453..4832
/gene="pmbA"
/locus_tag="Pro0996"
3453..4832
/gene="pmbA"
/locus_tag="Pro0996"
/note="pmbA"
/codon_start=1
/transl_table=11
/product="Predicted inactivated Zn-dependent protease"
/protein_id="AA000041.1"
/db_xref="GI:33237974"
/translation="MTMSNNBSLISEQNRLDKLTLSNEIKTKKMDLGASIGKTSVQ
IDKGEPRQIKASQSKSITIRVWKNENAVGITSIDSEAGLRKALNSAFSASFANKL
ESPNFSKLSQNALPNLQRPISQSLGVKLLAYIEASLKLKDNISIKIPYGALES
KYRKYVINSNGAFRHESTQASLYAPAAEBGRKXRSRSGSLKLAAYADINDEGIN
EAAKTLRHLNVQPIKTNKYLTCRPEALDILTSPNLYNARSITIDGSLSKNSIG
EQTSSNINISDEGLHKDPGACTPGEGETPQNTITLIDSGILKNIISHBAYATRTGC
NPLGHGIGIAKIVSPDPVLTITKGTETAKYPLDQPLQHTKEEFVLENHHLAAGIKS

SOGFSLEPFGMLVKNBQIESATTAAGDIKLNKIIQIEDNOSTHOGVPHIIV
 DELAVTGA
 /gene="fmc"
 /locus_tag="Pro00997"
 4829..5848
 /locus_tag="Pro00997"
 4829..5848
 /gene="fmc"
 /locus_tag="Pro00997"
 /codon_start=1
 /transl_table=11
 /product="Methionyl-tRNA formyltransferase"
 /protein_id="AAQ00042.1"
 /db_xref="GI:33237975"
 /translation="MNIITFMGTPIFCVPILEKLLKSNHNVAVITOPDRGRGNKVL
 PSPIKOKALEENLPYTPFNITKEKDOIARKYADIPVYVAGQILPKSVLKPRY
 GCWNTHALPRWRGAFTQWILSGSETEGVGJLWABEGLDTAVALLEKLLKULKE
 NAEQSQRLKDLSCNLIIEGILEKXAEQNSKLNLTQEDINTEYARLKLIE
 DYLIINMGYNHKKOILGLYPCYVINSKRLKLSIPINNKYDEDFRYSIDIL
 KSGIINNNGLIIDIEINIGIIVQTKVPLITVKLEKESQNALIQDLTINSG
 NKIG"
 complement(5849..6571)
 /gene="terc"
 /locus_tag="Pro00998"
 complement(5849..6571)
 /gene="terc"
 /locus_tag="Pro00998"
 /note="possibly involved in tellurium resistance"
 /codon_start=1
 /transl_table=11
 /product="Conserved membrane protein Terc"
 /protein_id="AAQ00043.1"
 /db_xref="GI:33237976"
 /translation="MDTASLDYLPFLNGIDQWSEFATLLPFLVLELLISADNAVAL
 ASITRDLSTLEOKSLNIGIATLPIILLIANTIIKYPITQIAASVLSIVT
 KPTINPNDESSISNNKSPFKITPELLATPLADLPAVAIVASQIILVITG
 ALIGYIARFTRADLFIRWDLFVQLENAGTAVAVAKILFEYTFNIVIPRITPI
 LLGFVFLMGFSKRVPLNSENE"
 6665..7375
 /locus_tag="Pro00999"
 6665..7375
 /locus_tag="Pro00999"
 /note="conserved in cyanobacteria"
 /codon_start=1
 /transl_table=11
 /product="Uncharacterized protein"
 /protein_id="AAQ00044.1"
 /db_xref="GI:33237977"
 /translation="MOINSELCHVDDEKIIIVITALENOIQSALGSGSNVHAES
 AIOOLITLKEVOKIOSINSNEAKLSVONSSTKEIEPKSKENSKISKDNVYD
 IYDCKVPMDSERISKDIERLGMKREBENPKITIGOSNRNPKSKYEDLPFI
 ELQCKRKGDGMKININYLEKULISENDILRLSNMSTAREYFVDFKRVNSSEL
 GIYELKFSMLGKELKS"
 7440..10952
 /gene="mfd"
 /locus_tag="Pro1000"
 7440..10952
 /gene="mfd"
 /locus_tag="Pro1000"
 /note="superfamily II helicase"
 /codon_start=1
 /transl_table=11
 /product="Transcription-repair coupling factor"
 /protein_id="AAQ00045.1"
 /db_xref="GI:33237978"
 /translation="WALKSLITNVESALTEIYKRVKCKKLSINGGTSAAITTS
 SLAKSNRPVIVPTLEASRMSIMQILGMNVLLPSSSEVSPYSISIKSEIWA
 OLVSLELANDYRDPALICTERALQPHLRPRLMENSYSIKGELMDLNSRWA
 IKIYEKTTQEGTGRGDIIDIYAVNNELPIRIELFDGNDIXREPDPINQRL
 DEINNRISPVINKITPNLNNELKKNKDIKHEIEGTIEELLEETQAKDSILMY
 IHKSLICIDKQCVSHENWNIHRTPELRMDNTKDNNSINTLNNDKQINS
 ILDEPKTIEKNDISINQDSPHNSKITYTYNQSKISELIERQKKKYSIWLK
 AOPSASALBSEHECTIKFVQNNLDISHINLIFKIPVSKLHGNLDIGVDSQK
 VLLTDKELFGQIVSTGTRKRSSSRSVNPKKSGDGVVHRNHGKFLKLEK

FIVNESHEDYLLVQXSDGTBVAADQSLGRVNSDPSKSPINKLGNWTAKERA
 KKSIRVADILIKIAERSNSOGSEFPDQGMORELEBAFYEATPDQIKAVSEIKSD
 MEKSPMDRLVGVDFGKTEVALPALKKAITAGQVAIAPVTVLAQOHWRTLIDRF

Query Match 28.6%; Score 161.8; DB 1; Length 302399;
 Best Local Similarity 60.3%; Pred. No. 2.1e-34;
 Matches 302; Conservative 0; Mismatches 197; Indels 2; Gaps 2;

QY 48 CATGAAACACACCCCTCTGTTTGTAGTGAAGATGAAGCCGAGTCTAACCAGATTCG 107
 DB 242767 CATGAGACACCTCTTCACTAGTCTAGTAGAGATCAATCAGGTCTTATGTAATTC 242826
 QY 108 CGGACTATTTGCCCGCGGTGTATTAACTAGAGCTTGGCGGTGGGGTGGCGGAGACA 167
 DB 242827 AGGCTATTTGCTCCGAGAGATTCACATCAACAGCCTTGTGTGGGCTGTGAGAGC 242886
 QY 168 GGGGACGTTTCCCGCATCACCATGCTGTCGCGGGAGTGAACACCATCAACAAC 227
 DB 242887 TGAAGCATTTCTCGCTTAACATGCTGTGAAGAGATCATCAACATCCACAT 242946
 QY 228 GACCAAGCACTCTCAAGTGTGTAAAGTAACTTAAGTACAGCATCACCGAACTCC 287
 DB 242947 GACCAAGCACTTGTATTAATTAATCAATGATTAAGTATGATCTAACAACATTTCC 243006
 QY 288 CTGTGTGAAGAGAAATTGATGCTGTGAAGGTGAGCGCCATGCCCCCTAACCGAGCGA 347
 DB 243007 TGCCGTAGAGAGAAATTATCTTTAAAGTATCTGCAACGAATGAAGAAAGAGCAAG 243066
 QY 348 AGTGATGAGCTAGCCAGCTATTCGGGCGCCGATTCGATCTCGAAGACACCGT 407
 DB 243067 CATCTTGATTTAGTTCAAGTTTAAAGCAAAAGTGTAGTATGAGAGATGCTCT 243126
 QY 408 CACCATCGAA-TGTTGGGGGACCCGGTAAATGTAGCAATCTCCAGATGTTG-GCCA 465
 DB 243127 TACATTAGAGAGTTGATAGAGATTCACGAGGAACTGGTAGCCCTAGAAAAGCTTTGAAGC 243186
 QY 466 AGTTGCATTAAAGAGGTGGCTCGAACCGGCAAAATGCTTTGTGTGGCGGAATCCGGGCT 525
 DB 243187 TTAGTCATTCTTGAGTATGAGCAAGAACTGGAAGATTCACCTGAAAGGCGCTCAGGTGT 243246
 QY 526 CATACGAAATATCTGAATC 546
 DB 243247 TAATACGAGATGTTGAATC 243267

RESULT 14
 AF041468/c 121524 bp DNA circular PLN 03-MAR-1999
 LOCUS
 DEFINITION
 ACCESSION
 X62348 X14171 X62349 X51511 X14504 X52158 X52912 X56806 W76547
 X62348 Z21576 U81044 AF063017
 AF041468.1 GI:3602932
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 chloroplast *Gulliardia theta*
Gulliardia theta
 Eukaryota; Cryptophyta; Cryptomonadaceae; *Gulliardia*.
 REFERENCE
 1 (bases 47701 to 48415)
 Douglas, S.B. and Durnford, D.G.
 The small subunit of ribulose-1,5-bisphosphate carboxylase is
 plastid-encoded in the chloroplast c-containing alga *Cryptomonas*
phi.
 JOURNAL
 Plant Mol. Biol. 13 (1), 13-20 (1989)
 MEDLINE
 9357429
 PUBMED
 2562756
 REFERENCE
 2 (bases 18535 to 19351)
 Douglas, S.B. and Durnford, D.G.
 Sequence analysis of the plastid rDNA spacer region of the
 chloroplast c-containing alga *Cryptomonas phi*.
 JOURNAL
 DNA Seq. 1 (1), 55-62 (1990)
 MEDLINE
 92119320
 PUBMED
 2132959
 REFERENCE
 3 (bases 43739 to 44938)
 Douglas, S.B. and Durnford, D.G.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE

Nucleotide sequence of the genes for ribosomal protein S4 and tRNA(Arg) from the chlorophyll c-containing alga *Cryptomonas phi*
90245597
2336372
4 (bases 34539 to 35380)
Reith,M. and Douglas,S.
Localization of beta-phycoerythrin to the thylakoid lumen of *Cryptomonas phi* does not involve a signal peptide
Plant Mol. Biol. 15 (4), 585-592 (1990)
91338697
2102376
5 (bases 45872 to 47981)
Douglas,S.E., Durnford,D.G. and Morden,C.W.
Nucleotide sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from the chlorophyll c-containing Alga *Cryptomonas F*: evidence supporting the polyphyletic origin of Plastids
J. Phycol. 26, 500-508 (1990)
Douglas,S.E.
Unusual organization of a ribosomal protein operon in the plastid genome of *Cryptomonas phi*: evolutionary considerations
Curr. Genet. 19 (4), 289-294 (1991)
91330343
1868578
7 (bases 40675 to 42376)
Douglas,S.E. and Turner,S.
Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor
J. Mol. Evol. 33 (3), 267-273 (1991)
92099311
1757997
8 (bases 96123 to 98906)
Wang,S.H. and Liu,X.Q.
The plastid genome of *Cryptomonas phi* encodes an hept0-like protein, a histone-like protein, and an acyl carrier protein
Proc. Natl. Acad. Sci. U.S.A. 88 (23), 10783-10787 (1991)
92073372
1961745
9 (bases 106789 to 108216)
Douglas,S.E.
A secy homologue is found in the plastid genome of *Cryptomonas phi*
FEBS Lett. 298 (1), 93-96 (1992)
92183838
1544427
10 (bases 42198 to 44153)
Douglas,S.E. and Reith,M.E.
A bohi homolog, encoding a subunit of Mg chelatase, is located on the plastid genomes of red and *Cryptomonad* algae
J. Mar. Biotechnol. 1 135-141 (1993)
Douglas,S.E. and Murphy,C.A.
Structural, transcriptional and phylogenetic analyses of the atpB gene cluster from the plastid of *Cryptomonas F* (Cryptophyceae)
J. Phycol. 30, 329-340 (1994)
12 (bases 98901 to 114602)
Wang,S.L., Liu,X.Q. and Douglas,S.E.
The large ribosomal protein gene cluster of a *Cryptomonad* plastid: gene organization, sequence and evolutionary implications
Biochem. Mol. Biol. Int. 41 (5), 1035-1044 (1997)
97283757
9137835
13 (bases 61067 to 68605)
Leitsch,C.E.W., Kowalik,K.V. and Douglas,S.E.
The alpha gene cluster of a *Cryptomonad*, *Gulliardia theta*: A piece in the puzzle of chloroplast genome development
J. Phycol. (1998) In press
14 (bases 1 to 121524)
Douglas,S.E. and Penny,S.L.
The plastid genome of the cryptophyte alga, *Gulliardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

J. Mol. Evol. 48 (2), 236-244 (1999)
99128221
9929392
15 (bases 1 to 121524)
Douglas,S.E.
Direct Submission
Submitted (08-JUN-1998) Institute for Marine Biosciences, National Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada
On or before Sep 15, 1998 this sequence version replaced gi:11396, gi:11297, gi:18103, gi:18281, gi:11383, gi:11407, gi:12539, gi:136730, gi:11300, gi:398949, gi:2661180.
Location/Qualifiers
1. 121524
/organism="Gulliardia theta"
/organism="plastid:chloroplast"
/mol_type="genomic DNA"
/db_xref="taxon:55229"
/note="plastid"
complement(5..86)
/gene="trnY(gua)"
complement(5..86)
/gene="trnY(gua)"
/product="tRNA-Tyr"
/note="codon recognized: UAC"
/anticodon=(pos:50..52,aa:Tyr)
complement(181..573)
/gene="rpl19"
complement(181..573)
/gene="rpl19"
complement(181..573)
/product="ribosomal protein l19"
/protein_id="AAC35594.1"
/db_xref="GI:3602934"
/translation="METILLTYEFVNKSKNLLIQLVESHKNTLPQLQVDTVRIGVSTIEGSKRIQATEGVITISKNGINLTITIRVVOGIGIERVYLLHSPKLSIELRKSRIKRAKLYVIRTRGKATRLQQRN"
729..3188
/gene="clpC"
729..3188
/gene="clpC"
/codon_start=1
/product="Clp protease ATP binding subunit"
/protein_id="AAC35595.1"
/db_xref="GI:3602934"
/translation="MFERFPEKAIKVIMLAQEARSLGHNVGTEDLIGLIGEGTGI AAVKLSMGVNLDAVEVEKTIIRGSGPVAVIRPTPRAKRILELSLEBARLGHNY IGHVILLIGIRGEGVAAVLENALDLTKVTQVIRLIGTDAVSATKQTKTP TIEFGSNLQKABEGLDLPVIGRQKEIEKVIQILDRITKNPILIGEPVGATAIAE GLAQRLNNRPVPLIEDKRVTLIDIGLVAQKYRGEFEBRLKIIIDELRVANNVILV IDEVHTLIGAGAEAGAIIDANILIKPALAGEMQICATLIEERKHI EKOSALERFO PVWGESVSEETIEILYLRDRKEKHKVIISBALSAAKAPADQYIADRPDKXID LIDEGASRVAKNSOLPPABRELDELRIKQKDRVAVSODETGCGLRDEPMETKA GTIAIHSKKKDEENKEVSVTEEDIIQIVANGTIPNNKTRSESKILQNEETLH GRIICGDEAVNAVSAKIRRAVATKZPNKPNPFIISGTGFGKTELTAALAFVFGS ERMVFLDMSSEVIERHTVSLISPGVYVNEGQLTSSVRRRPYTVLAFBIEFGH PDVFNNLLQILEDGLTDSKRTVDPKNTLLITSNVGSKVEKGGGGLGFDISEQT ESQVGRILKLVNELKQYRPERPLANLDEIIVFQULTKDEVGIIAIMKEVPTRISE KGIQLEVTARFKTHLINEGNPNPYGARPIPRAVMRLLEDLTLSSEFLAEXIKEDTVAV DVDDDKKVTALGEXKELVAN"
3316..3388
/gene="trnF(gaa)"
3316..3388
/gene="trnF(gaa)"
/product="tRNA-Phe"
/note="codon recognized: UUC"
/anticodon=(pos:3349..3351,aa:Phe)
complement(3402..5201)
/gene="dnaB"
complement(3402..5201)
/gene="dnaB"
/codon_start=1
/transl_except=(pos:complement(5199..5201),aa:Met)

	Query Match	Best Local Similarity	Matches	298; Conservative	28.2%; Score 159.6; DB 8; Length 121524; Prid. No. 8e-34; Mismatches 194; Indels 2; Gaps 2;
QY	49	ATGAAACACACCCCTCTCTGTTTAGTGAAGTAAGACCGGAGTGTAAACCGCATTTGCC	108		
DB	50006	ATGAAACATACATTATTCAGTTTGTTGTTGAAGATGAAATCGTGTTTAACTCGTATGCT	49947		
QY	109	GGACTATTGCGCCGCGTGGTTTAAACATGAGAAGCTTGCGGCTGGCGGCGGAACG	168		
DB	49946	GGTTATTCCCTCGTCGTGTTCAATTCGAAAGTTTAGACATTGGAACCACTGAAAA	49887		
QY	169	GGGAGCGTTCCCGCATCACATGCGTGGTCCGCGGAGTAGAACAACCATCGAACACTG	228		
DB	49886	TTAGAAATTCACGTATTCTATGTTGTTTCCATCGAGATGAAGACTATTGAGCAATTA	49827		
QY	229	ACCAAGCAACTCTACAAAGTTGGTTAACGTAATTAAAGTACAGACATCACCGAACTCCC	288		
DB	49826	ACAAAACAACCTTATAAGCTTGTTAATATATTAAAGTTGMAAGTATTCGAACTCACT	49767		
QY	289	TGTGTGGAAGGAATTGATGCTGCGTGAAGTGAAGGCCAATGCCCTTAACGAGCGGAA	348		
DB	49766	ACGGTTAGCGAGAACATCATGTTGATTAAATTCAGATTTCAACATTTGAGAGACTGAA	49707		
QY	349	GTCATTAGCTAGCCACAGTATTTCCGGGCCCGCATTTGTGATATCTCCAGACACCGTC	408		
DB	49706	GCATTAAATTAATTAATATATTTTCGTGCTAAGGTGTGAACCTATCTGAGATTTT	49647		
QY	409	ACCATCGAA-TGCTGGGGGACCCGGGGTAAATGGTAGCAATCCTCCAGATGTTGGCCAG	467		
DB	49646	ATAATTAAAGTTACAGGTGATCCGTGTAATTTGTGCCATTAGAGCAAAATTTGTAAAA	49587		
QY	468	TTGGC-ATTAAAGAGTGTGCTCGAAACGGCAAAATGCTTTGGTGGCGGATCCGGCGTC	526		
DB	49586	TTTACTATATCTGAGATAGTAAGACTCGAAAAATTTCTCTTGTAGCGCAATCTAATGTT	49522		
QY	527	AATACGGAATATCTT 540			
DB	49526	AATACCGGTTCTCT 49513			

	DEFINITION	Streptomyces cinnamonensis acetylhydroxy acid synthase small subunit (11vN) gene, complete cds.
	VERSION	AF175526
	KEYWORDS	AP175526.1 GI:5733115
	SOURCE	Streptomyces cinnamonensis
	ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycetaceae; Streptomyces.
	REFERENCE	Kopecky,J., Janata,J., Pospisil,S., Felsberg,J. and Spizek,J. Mutations in two distinct regions of acetylactate synthase regulatory subunit from Streptomyces cinnamonensis result in the lack of sensitivity to end-product inhibition Biochem. Biophys. Res. Commun. 266 (1), 162-166 (1999)
	JOURNAL MEDLINE	20050069
	PUBMED	10581183
	REFERENCE	2 (bases 1 to 593) Kopecky,J., Janata,J., Pospisil,S., Felsberg,J. and Spizek,J. Direct Submission Submitted (05-AUG-1999) Genetics and Physiology of Actinomycetes, Institute of Microbiology, Videnska 1083, Praha 4 CZ-14220, Czech Republic
	FEATURES	Location/Qualifiers
	source	1..593 /organism="Streptomyces cinnamonensis" /mol_type="genomic DNA" /strain="C-100-5" /db_xref="taxon:1900" 30..557 /gene="11vN" 30..557 /gene="11vN" /codon_start=1 /transl_table=1 /product="acetylhydroxy acid synthase small subunit" /protein_id="AAD49432.1" /db_xref="GI:5733116" /translation="MSTRKLTLVLENKPGLARLTFSRGNIDSLAVGTEHPEDISRTIVNVNEDPLBOYTKQINKLVNLKIVLEPPAAVARLVLAKVADNETRSQIVEIVOLPRAKTVDVSPEAATVIETATGSDDKLEMLKMLBEPGRVELVQSITAIIGRSSRIIDRSIRALDRSA"
	gene	
	CDS	
	ORIGIN	
	Query Match	27.0%; Score 152.4; DB 1; Length 593;
	Best Local Similarity	59.6%; Pred. No. 5,7e-32;
	Matches	292; Conservative 0; Mismatches 196; Indels 2; Gaps 2;
Dy	CCATGAACACACCCTCTGTGTTTAACTTGTAAGATGAAGCCGAGTGCTAACCGCATGG	106
Dy	34 CCACCAACACACCGCTTCGCTCCTGTCGAGAACCAAGCCGCGCTCTCCCGCGATCA	93
Dy	107 CCGACTATTGCCCGCGCTGTTTAACTTAGAGCTTGCGGTGGGGTCGGCGAGAC	166
Dy	94 CGCGCTGTTCTGGCGCGGCTTCAACATCGACTCGCTCGGTGGCGTCAACGAGC	153
Dy	167 AGGGGAGCGTTCGCATCACCATGGTGGTGCGGGGAGTAGAACAACATCGAACAC	226
Dy	154 ACCCGGATCTCCCGATCAACATCGTGGAGAACGTGAGAGACCTGCGCTCGAACGG	213
Dy	227 TGACCAAGCAACTTCAAGTTGGTTAACGTAACTTAAAGTACAGAGACTACACCAACTC	286
Dy	214 TGACCAACCACTCAACAAAGCTGTCCAACGCTTCMAATGATGTCAATGAGAGCCGGCG	273
Dy	287 CCTGTGTGAAGAAGGAATTGATGCTGGTGAAGTGAAGGCCAATGACCTTAACGAGCGG	346
Dy	274 CCGCGTGGCCCGGAGCTGCTCTGGGGAAGTCCGGCGCGAACAAGAACCCGCTCC	333
Dy	347 AAGTATTAGCTAGCCCCAGGTATTCCGGGCCCGCGCATTTGGATATTCTCCGAAGACCTG	406
Dy	334 AGATGCTGAGATCGTCCAGCTGTTCGCGGCAAGACCCGTGACGCTTCGCCGAGGCGG	393
Dy	407 TCACCATGAATGGTGGGGGACCGG-GGTAATAATGTGCAATCTCCAGATGTTGG-CC	464

Db	394	TCACCATCGAGGCGACCGGCAAGCTGAGGCGCATGCTCAAGATGCTGAGC	453
QY	465	AAATTGGCATTAAAGAGGTGCTCGAAGCGGCAAAATTTGTTGTGCGGAAATCCGCG	524
Db	454	CCTTCGGCATCAAGAGTGTGCAAGTCCGCGACCATCGGATCGGCGCGGCTCCCGCT	513
QY	525	TCAATACGGA	534
Db	514	CCATCAAGGA	523

Search completed: July 25, 2004, 10:17:21
 Job time : 2273.95 secs

